

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.4 p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 12, 2003, 10:35:25 ; Search time 5814 Seconds
(without alignments)
7273.199 Million cell updates/sec

Title: US-09-759-130B-381
Perfect score: 8138
Sequence: 1 MMLPQNSWHDGRCCHQN.....CEDASDTSLLGVLPASEATK 1453

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus.p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09759130/runat_06052003_123554_23353/app_query.fasta_1.1607
-DB=GenEmbl -QFWT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosom62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=2000000000
-USER=US09759130.cgn_1.1.3897 @runat_06052003_123554_23353 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_or.*
21: em_ov.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_hgt_hum.*
31: em_hgt_inv.*
32: em_hgt_other.*
33: em_hgt_mus.*
34: em_hgt_pin.*
35: em_hgt_rod.*
36: em_hgt_mam.*
37: em_hgt_vrt.*
38: em_sy.*
39: em_hgtgo_hum.*
40: em_hgtgo_mus.*
41: em_hgtgo_other.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	8109	99.6	4579	9	AF264014 Homo sapi
2	6107.5	75.0	3670	6	AX133829 Sequence
3	3779	46.4	2249	9	AK058038 Homo sapi
4	3441.5	42.3	3786	9	HSML30AC1
5	3429	42.1	3703	9	HSML30A
6	3429	42.1	4950	9	HSML30AC2
7	3396.5	41.7	3802	9	HSML30AE
8	3377	41.5	3707	6	AX281768
9	3272	40.2	4405	10	AF274883
10	3082	37.9	4475	4	BBWC11MR
11	2990	36.7	7656	9	HS243224
12	2987	36.7	7683	9	HS243212
13	2982	36.6	7667	9	HS297935
14	2980	36.6	7686	9	AF159456
15	2783	34.2	202252	2	AC131207
16	2447.5	30.1	7006	3	AF064259
17	2435	29.9	5802	6	A90822
18	2435	29.9	5802	6	AR193170
19	2435	29.9	5802	6	HSDBMT1
20	2246.5	27.6	6323	3	GFY14953
21	2161	26.6	5108	4	AF043112
22	1963.5	24.1	6652	10	MMU37438
23	1856.5	22.8	6290	10	AB005909
24	1594.5	19.6	3769	3	AF228827
25	1586	19.5	2305	4	AF400115
26	1465	18.0	170882	2	AC107839
27	1442	17.7	3431	3	AF228826
28	1399	17.2	2423	4	SSSRP5
29	1341	16.5	1665	4	SSSRP3
30	1305.5	16.0	177768	2	AC131206
31	1304	16.0	3927	3	AF228824
32	1302	16.0	205331	2	AC124470
33	1277	15.7	3332	3	AF228825
34	1263	15.5	170882	2	AC107839
35	1205	14.8	2165	4	AF400122
36	1204	14.8	1316	4	SSSRP2
37	1176	14.5	161644	9	AL161645
38	1142.5	14.0	4344	6	AX402036
39	1142.5	14.0	4344	10	RNU32681
40	1142.5	14.0	4360	6	172544
41	1085.5	13.3	1519	4	SSSRP4
42	1078.5	13.3	1800	4	S78981
43	1041.5	12.8	1654	5	AB051832
44	968.5	11.9	2806	9	HS243212
45	916	11.3	192430	2	AC079542

ALIGNMENTS

RESULT 1

Db 326 GGACACCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCTCTGTTATGGAAAT 385
Qy 121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140
Db 386 GAGTCAGCTCTCTGGGAATGTCACACCCGGGATGGGAAGCCATAACTGTTATCATGGA 445
Qy 141 GluAspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGly 160
Db 446 GAAGATGTTGGTGTGAACCTTACCGTGAAGCCAATCTGGGTTTGAGGCTAGTGGATGGA 505
Qy 161 AsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCys 180
Db 506 AACAACTCCCTGTTACGGAGAGTGGAGGTGAAATTCACAAAAGTGGGAACATATATGT 565
Qy 181 AspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer 200
Db 566 GATGATGGGTGGAACTTGAATACCTGTCGCGTGGTGTGCAGGCNACTAGGATGTCATCT 625
Qy 201 SerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAsp 220
Db 626 TCTTTTATTTCTCTGGAGTTGTTAATAGCCCTGCTGATTTGCCGCCCAATTTGGCTGGAT 685
Qy 221 AspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGly 240
Db 686 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTCGACACATCGTGGATGGGA 745
Qy 241 AsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeu 260
Db 746 AYTCACTGCTGCAGTCAACAATGAGGATGTCACATTAATCTGTTATGATAGTAGTGATCTT 805
Qy 261 GluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGln 280
Db 806 GAACAAAGGCTGTGAGTGGAACTAACCGCTGATGGGAGAGTAGAGCTGAAATCCAA 865
Qy 281 GlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAspValValCys 300
Db 866 GGAAGGTGGGGAGCGTATGCCACCAATAGTGAACATGCTGAGCTGATGCTGATGC 925
Qy 301 LysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGly 320
Db 926 AAGCAGTTGGAGTGGAAACCGCACTTCACTTCGCTGGCTTGCCTCAATTTGCAGTCAGG 985
Qy 321 SerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAsp 340
Db 986 TCTGATCTGTATGGCTGTATGGTGTCTCTGCTCCGGTAATGAATCTTTCTTTTGGAC 1045
Qy 341 CysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValle 360
Db 1046 TGCAGACATTCGGGAACCGCAATTTTGAAGTGTCTTCAATCAAAACGATGCTCTGTATC 1105
Qy 361 CysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGly 380
Db 1106 TGCTCAGATGGACAGATTTGGAACTCCGACTAGCAGATGGAAGTAACAATTTGTCAGG 1165
Qy 381 ArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsn 400
Db 1166 AGACTAGAGTGAACAATTCATGAACACTGTTGGACAATATGTGACCAAGAACTGGAAGAAT 1225
Qy 401 GluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArg 420
Db 1226 GAACAAGCCCTCTGTGGTTGTTAAGCAGCTAGGATGTCGGTTCAGCGTCTTTGGCAGTCGT 1285
Qy 421 ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly 440
Db 1286 CGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGACATGGG 1345
Qy 441 AsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArg 460
Db 1346 AATGAGTCACTCTCTGGGACTGCACATATGATGGAANAAGCAAGCAACATGCTTCCGA 1405
Qy 461 ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGly 480
Db 1406 AGATCAGATGCTGGAGTAATTTGTTCTGTATAAGGCAGATCTGGACCTTAAGGCTTGTGGG 1465

Qy 481 AlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrVal 500
Db 1466 SCTCATAGCCCTCTTATGGAGATTGGAGGTGAATACCAAGAGAGTGGGGACTGTG 1525
Qy 501 CysHisAspArgTrpSerThrArgAsnAlaValValCysLysGlnLeuGlyCysGly 520
Db 1526 TGTCATCAGAGATGGAGCACAAGGAATGCACTGCTGTGTAAACAATTTGGGATGGA 1585
Qy 521 LysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeu 540
Db 1586 AAGCCTATGATGTTGTTGGTATGACCTATTTAAAGAAGCATCAGGACCTATTTTGGCTG 1645
Qy 541 AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp 560
Db 1646 GATGACCTTCTTGCATTTGGAATGACTCAATATCTGGGACTGTGAACACAGTGGATG 1705
Qy 561 GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr 580
Db 1706 GGAAGCATTAATTTGTTACACAGAGAGATGTTGATTTGTAACCTGCTCAGGTGATGCA 1765
Qy 581 TrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPhe 600
Db 1766 TGGGGCTGAGGCTGGTGGCGCGCAGCAACCGCTGCTCGGAAGACTGGAGGTGTACTTT 1825
Qy 601 GlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValVal 620
Db 1826 CAAGGACGGTGGGCGACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGGTG 1885
Qy 621 CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr 640
Db 1886 TGTAGCCAGCTGGACTGCCATCTTCTATCATTTGGCTGGTCTGGGAAACGCTTCTACA 1945
Qy 641 GlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrp 660
Db 1946 GGATATGGAAAAATTTGGCTCGATGATGTTTCTCTGATGGAGATGACTCAGATCTCTCG 2005
Qy 661 SerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyVal 680
Db 2006 TCATGACAGAACACTGGGTGGGAAATAATGACTGCAGTGCACATGAAGATGTTGGAGTG 2065
Qy 681 IleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAla 700
Db 2066 ATCTGTCTCTGATGATCGGATATGAGCTGAGGCTGTTGGTGGAAAGCAGCAGGTGCT 2125
Qy 701 GlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGly 720
Db 2126 GGAAAAATTTGAGGTGAATGTCACAGGTGGCGTGGGAATTTCTGTGTCTAATGGCTGGGA 2185
Qy 721 MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer 740
Db 2186 ATGAACATTTGCTGAAGTGTGTTGGAGCAACTTGAATGTGGGTCTGCAATCAGGCTCTCC 2245
Qy 741 ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr 760
Db 2246 AGAGAGCTCATTTACAGAAAGAACATTAACATCTTAATGTGCAATTTCTGGCTGCACT 2305
Qy 761 GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis 780
Db 2306 GGAGGGAAAGCCCTCTCTCGGATTTATACGATGGAGTGGAGTGGAAACAGACTCCGTGTCT 2365
Qy 781 LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyVal 800
Db 2366 TTAATATGAAGAACAGTTGATCTGCTACGCCACAGCCAGCCAGGCTGGTGGAGCT 2425
Qy 801 AspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys 820
Db 2426 GATATGCCCTGCTCTCGGACGTGTTGAAGTGAACATGCAGACACATGCGCTCTGTCTGT 2485
Qy 821 AspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp 840
Db 2486 GATTCTGATTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2545

QY 841 AlarileSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu 860
DB 2546 GCCATATCTCTTCTGTGGAGATCACTTGGAAAAGGAATGCTCAACTTGGCCGAA 2605
QY 861 LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro 880
DB 2606 AAGTCCAGTGTGAAGGAGTGAACCTCACCTTGCAATATGCCCATTTGTTCAACATCCG 2665
QY 881 GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspVal 900
DB 2666 GAAGACACTTGTATCCACAGCAGAAAGTGGAGTTGCTCTTCCGATATACAGATGTC 2725
QY 901 ArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHis 920
DB 2726 CGACTTGTGAATGGCAATCCACAGTGTGACGGCAAGTGGAGATCAACGTCCTTGGACAC 2785
QY 921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940
DB 2786 TGGGGCTCACTGTGTGACACCACTGGGACCCAGAAAGATGCCCGTGTCTATGACAGACAG 2845
QY 941 LeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerVal 960
DB 2846 CTCAGCTGTGGGACGCTCTCTACACACAGGAGGAAATATATATGGAGAAAGTGT 2905
QY 961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGln 980
DB 2906 CGTGTGTGGGACACAGGTTTCAATGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAA 2965
QY 981 MetThrValLeuGlyAlaProCysIleHisGlyAsnThrValSerValIleCysThr 1000
DB 2966 ATGACAGTCTTGGAGACCTCCCTGTATPCCATGGAATACTGTCTCTGATCTGCACA 3025
QY 1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSer 1020
DB 3026 GGAAGCCTGACCCAGCAGCTTTCCATGCTCGCAATATGATATGACCCATATTTGCT 3085
QY 1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgValAsp 1040
DB 3086 GCAGTTCACAGAGGCGCTGTGATCTCTTAGAGCAAAACGGCTCCGCCCTAGTGGAT 3145
QY 1041 GlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIle 1060
DB 3146 GGGGACAGCGCTGTCCCGGAGAGTAGAGATCTATCACAGCGCTTCGTGGGACCATC 3205
QY 1061 CysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080
DB 3206 TGTGATGACGGCTGGGACCTGAGCGATGCCAGTGTGTCTCAAAAGCTGGGCTGTGA 3265
QY 1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeu 1100
DB 3266 GTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGTCAAGGCCCATCTGGCTG 3325
QY 1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120
DB 3326 GATGACCTGAACTGCACAGGAATGGAGTCCCACTTGTGCGAGCGCTTCCCGCGCTGG 3385
QY 1121 GlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAla 1140
DB 3386 GGGCAGCACCACTGCAGGCACAAGGAGGACGAGGGTCTATCTGCTCAGAAATTCACAGCC 3445
QY 1141 LeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyr 1160
DB 3446 TTGAGGCTCTACAGTGAACCTGAACAGAGAGTGTGCGGAGATGGAGTCTCTAT 3505
QY 1161 AsnGlyThrTrpGlySerValGlyArgAsnIleThrThrAlaIleAlaGlyIleVal 1180
DB 3506 AACGGACCTGGGCGACGCTCGCAGGAGGAAACATCACACAGCATAGCAGCAATGTG 3565
QY 1181 CysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThr 1200
DB 3566 TGCAGGCGAGTGGGCTGTGGGGAGATGGAGTTGTGAGCCCTCCGCCCTTTATCTAAGACA 3625
QY 1201 GlySerGlyPheMetTrpValAspAspIleClnCysProLysThrHisIleSerIleTrp 1220

DB 3626 GGCTGCTTTCATGTGGTGGATGACATTAGTGTCTCTAAACGCATATCTCCATATGG 3685
QY 1221 GlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIle 1240
DB 3686 CAGTGCCTGTCTCCCATGCGGAGCGAAGATCTCCAGCACACAGAGAGACCTGGATC 3745
QY 1241 ThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260
DB 3746 ACATGTGAAGATAGAATAAGAGTCCGTGGAGAGACACCCAGTGTCTGGGAGAGTGGAG 3805
QY 1261 IleThrHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280
DB 3806 ATCTGGCAGCAGGCTCTCTGGGACACAGTGTGTGATGACTCTCTGGGACCTGGCCGAGCG 3865
QY 1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAlaSer 1300
DB 3866 GAAGTGTGTGTACAGACCTGGGCTGTGGCTCTCTCTGGCTGCCCTGAGGAGCGCTCG 3925
QY 1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320
DB 3926 TTTGCCGAGGAACTGGAACTATCTGTTGGATGACATGCGGTGCAAGGAAATGAGTCA 3985
QY 1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340
DB 3986 TTTCTATGGGACTGTACGCCAAACCCCTGGGACAGAGTGTCTGTGACACCAAGGAAGAT 4045
QY 1341 AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuLeuAlaSerSerGlyHisLeu 1360
DB 4046 GCTGCGGTGAGTGTCTGTGGACAGTCTGGAATCACTGAATGCTCTCAGGTGCTTTA 4105
QY 1361 AlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPheIleLeuPheLeu 1380
DB 4106 GCATTTATTTATCCAGTATCTTTGGGCTCTCTCCCGTCTGTGTTATTTCTATTCTC 4165
QY 1381 ThrTrpCysArgValGlnLysHisLeuProLeuArgValSerThrArgArgArg 1400
DB 4166 ACGTGTCCCGAGTTTCAGAAAACAAACATCTGCCCTCAGAGTTTCAACAGAGGAGG 4225
QY 1401 GlySerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspPro 1420
DB 4226 GTTCTCTCGAGGAGAAATTTATTCATGAGTGGAGACTGCTCCTCAGAGAGAGGCCCA 4285
QY 1421 HisGlyThrArgThrSerAspThrProAsnHisGlyCysGluAspAlaSerAspThr 1440
DB 4286 CATGGACACAGAACTCAGATGACACCCCAACCATGTTGTGAAGATGCTAGCGACACA 4345
QY 1441 SerLeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453
DB 4346 TCGCTGTGGGAGTCTTCTGCTCTGAAGCCACAAAA 4384

RESULT 2

AXI33829

LOCUS

Sequence 15 from Patent WO0119856.

ACCESSION AXI33829

VERSION AXI33829.1

KEYWORDS

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..3670

/organism="Homo sapiens"

/db_xref="taxon:9606"

3670 bp DNA linear PAT 15-MAY-2001

Qy	837	AsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeu	856
Db	1622	AACTGGGAGATGCCATATCTCTTCTGGGAGATCACTTTGGAAAGGAATGCTTA	1681
Qy	857	ThrTrpAlaGlyPheGlnCysGluGlySerGlnThrHisLeuAlaLeuCysProfile	876
Db	1682	ACTGGGCGGAAAAGTTCCAGTGTGAAGGAGTGAACACTCACCTTGCATTAATGCCCATTT	1741
Qy	877	ValGlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArg	896
Db	1742	GTTCACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGAGTTGCTCTTCCCGCA	1801
Qy	897	TyrThrAspValArgLeuValAsnGlyLysSerGlnCysaspGlyGlnValGluIleAsn	916
Db	1802	TATACAGATGTCGCACTTGTGAATGGCAAAATCCCAAGTGTACGGGCAAGTGGAGATCAAC	1861
Qy	917	ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal	936
Db	1862	GTGCTGGACACTGGGGCTCACTGTGTACACCCACTGGGACCCAGAGAATGCCGTGTT	1921
Qy	937	LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGly	956
Db	1922	CTATGCACAGACTCAGCTGTGGGACTGCTCTCTCAACAGAGAGAAATATATTGGA	1981
Qy	957	GluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu	976
Db	1982	GAAGAAGCTTCGTGTGTGGGACACAGTTTCATTCTTAGGGAATGAGTCACCTCTG	2041
Qy	977	AspAsnCysGlnMethrValLeuGlyAlaProProCysIleHisGlyAsnThrValSer	996
Db	2042	GATAACTGTCAAAATGACAGTTCTTGAGCACCTCCCTGTATCCATGGAAATACTGTCTCT	2101
Qy	997	ValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAsp	1016
Db	2102	GTGATCTGCACAGAAACCTTGACCCAGCCACATGTTCCATGCCTCGCAATGATCTGAC	2161
Qy	1017	ProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeu	1036
Db	2162	CCATATTGTCAGTTCCAGAGGCACTGCTTTGATCTGCTTAGAGGACAAACGGCTC	2221
Qy	1037	ArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPhe	1056
Db	2222	CGCTAGTGAATGGGACAGCCGCTGTGCGGGAGAGTAGAGATCTATCACAGCGGCTC	2281
Qy	1057	TrpGlyThrIleCysAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLys	1076
Db	2282	TGGGSCACCATCTGTGATGAGCGCTGGGACCTGAGCGATGCCACGTGCTGTCAAAAG	2341
Qy	1077	LeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGly	1096
Db	2342	CTGGGCTGTGGAGTGGCTTCAATGCCAGGCTCTGTCTACTTTGGGAGGGGTCAAGG	2401
Qy	1097	ProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysPro	1116
Db	2402	CCCATCTGGCTGGATGACCTGAATGCACAGGAATGAGTCCCACTGCTGGCAGTGCCT	2461
Qy	1117	SerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSer	1136
Db	2462	TCCCGCGCTGGGGGACAGCACTGCAGCACAGAGGACGAGGGGTCTATCTGTCTCA	2521
Qy	1137	GluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeu	1156
Db	2522	GAATTCACAGCCTTGAGGCTCTACAGTGAACCTGAACAGGAGGCTGTGCTGGGAGATTG	2581
Qy	1157	GluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIle	1176
Db	2582	GAAGTCTTCTATACCGGCACTGGGCGAGCGTCCGAGGAGGAACATCACCACAGCCATA	2641
Qy	1177	AlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaPro	1196
Db	2642	GCAGGCATTTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAAGCTGCCCT	2701
Qy	1197	LeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHis	1216

Db	2702	TTATCTAAGACAGGCTCTGGTTTCATGTGGTGGATGACATTCAGTCTCTAAACGCAT	2761
QY	1217	IleSerIleTrpGlnCysLeuSerAlaProTrpGluArgIleSerSerProAlaGlu	1236
Db	2762	ATCTCCATATGCAGTGCTGTCTCCCATGGGAGGAAATCTCCAGCCACGAGAA	2821
QY	1237	GluThrTrpIleThrCysGluAspArgIleArgValIleArgGlyGlyAspThrGluCysSer	1256
Db	2822	GAGACCTGGATCACATGTGAAGATAGAATAAGATGGCTGGGAGAGACACCGAGTGTCT	2881
QY	1257	GlyArgValGluIleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAsp	1276
Db	2882	GGAGAGAGTGGAGATCTGGCACCCAGGCTCTCTGGGCACAGTGTGTGACTCTCTGGGAC	2941
QY	1277	LeuAlaGluAlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeu	1296
Db	2942	CTGGCCGAGGCGAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTGTGGCTGGCTGGCCCTG	3001
QY	1297	ArgAspAlaSerPheGlyClnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLys	1316
Db	3002	AGGAGCGCTTCGTTTGGCCAGGGAAGCTGAACCACTCTGGTTGGATGACATCGCGTGCATA	3061
QY	1317	GlyAsnGluSerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGly	1336
Db	3062	GGAATGAGTCAATCTATGGGACTGTACGGCCAAACCTCGGGACAGAGTGAAGTGTGA	3121
QY	1337	HisLysGluAspAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSer	1356
Db	3122	CACAAGGAAGATCTCGGCTGAGGTGCTCTGGACAGTGCCTGAAATCACTGAATGCCTCC	3181
QY	1357	SerGlyHisLeuAlaLeuIleLeuSerIlePheGlyLeuLeuLeuValLeuPhe	1376
Db	3182	TCAGGTCAATTAGACATTAATTTATCCAGTACTTTTGGGCTCTCTCTCGTGTCTGTTT	3241
QY	1377	IleLeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSer	1396
Db	3242	ATTCTATTCTCACGTGGTGGCGAGTTCAGAAACAAACAAACATCTGCCCTCAGAGTTTCA	3301
QY	1397	ThrArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLys	1416
Db	3302	ACCAGAAGGAGGGTCTCTCGAGGAGAATATTTTCCATGAGATGGAGACCTGGCTCAAG	3361
QY	1417	ArgGluAspProHisGlyThrArgThrSerAspAspThrProAsnHisGlyCysGluAsp	1436
Db	3362	AGAGAGACCCACATGGGCAAGAACTCAGATGACACCCCAACCATGGTTGTGAAGAT	3421
QY	1437	AlaSerAspThrSerLeuLeuGlyValLeuProAlaSerGluAlaThrLys	1453
Db	3422	GCTAGCCACACATCGCTGTGGGAGTCTTCTGCTCTGAAGCCACAAAA	3472

DECEMBER 2

AK058038

LOCUS

DEFINITION

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

ACCESSION

VERSION

KEYWORDS

SOURCE

QY 1310 LeuAspMetArgCysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysPro 1329
 DB 1623 TTGGATGACATCGGTGCAAGAAATGAGTCAATTTCTATGGAGTGTACGCCAAACCC 1682
 QY 1330 TrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSer 1349
 DB 1683 TGGGACAGAGTGAAGTGGACACAGGAAGTGTGGCGTGGAGTGTCTGGACAGTCG 1742
 QY 1350 LeuLysSerLeuAsnAlaSerSerGlyHisLeuAlaLeuLeuLeuSerSerIlePheGly 1369
 DB 1743 CTGAATCACTGAATGCTCTCAGTCACTATTTAGCACATATTTATCAGTATCTTTGGG 1802
 QY 1370 LeuLeuLeuValLeuPheLeuLeuPheLeuTrpCysArgValGlnLysGlnLys 1389
 DB 1803 CTCTTCTCTGTTCTGTTATTTCTTCTACGTGGTGGCGAGTTCAGAACAAAA 1862
 QY 1390 HisLeuProLeuArgValSerThrArgArgArgGlySerLeuGluGluAsnLeuPheHis 1409
 DB 1863 CATCTGCCCTCAGAGTTTCAACCAAGAGGAGGTTCTCTCGCGGAGAAATTTATTCAT 1922
 QY 1410 GluMetGluThrCysLeuLysArgGluAspProHisGlyThrArgThrSerAspAspThr 1429
 DB 1923 GAGATGGAGACCTGCTCTAAGACAGAGACCCACATGGGACAGAACCTCAGATCACACC 1982
 QY 1430 ProAsnHisGlyCysGluAspAlaSerAspThrSerLeuLeuGlyValLeuProAlaSer 1449
 DB 1983 CCCAACCATGGTTGTGAAGATGCTAGCACACATCGCTTTGGGAGTCTTCTCTGCTCT 2042
 QY 1450 GluAlaThrLys 1453
 DB 2043 GAAGCCACAAAA 2054
 RESULT 4
 HSM130AC1 3786 bp mRNA linear PRI 22-NOV-1996
 LOCUS H.sapiens mRNA for M130 antigen cytoplasmic variant 1.
 DEFINITION 22969
 ACCESSION 22969.1 GI:312143
 VERSION 1
 KEYWORDS antigen; M130.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 MICKLEM,K.K.
 DIRECT SUBMISSION
 TITLE Submitted (14-JUN-1993) Kingsley K.J., Micklem, Nuffield Department
 of Pathology, University of Oxford, Level 1, Maternity Block, John
 Radcliffe, Hospital, Headington, Oxford, OX3 9DU, United Kingdom
 REFERENCE 2 (bases 1 to 3786)
 AUTHORS Law,S.K., Micklem,K.J., Shaw,J.M., Zhang,X.P., Dong,Y., Willis,A.C.
 and Mason,D.Y.
 TITLE A new macrophage differentiation antigen which is a member of the
 scavenger receptor superfamily
 JOURNAL Eur. J. Immunol. 23 (9), 2320-2325 (1993)
 MEDLINE 93380506
 PUBMED 8370408
 FEATURES Location/Qualifiers
 1..3786
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_type="Macrophage"
 102..3557
 /codon_start=1
 /product="M130 antigen cytoplasmic variant 1"
 /protein_id="CAA80542.1"
 /db_xref="GI:312144"
 /db_xref="SPTREMBL:Q07899"
 /translation="MVLLEDGSGADFRRHFNVLSPFTITVVLISACFVTSLSGTDK
 ELRLVDGKSGRVEYVQEVNWSMEAVSYICNLQGLCPATAKAPGWNSS
 AGSGRIWIDKSNESALMDCKHDGKHSNC"PHOODAGVTCSDGNSNLEMLRTRGG
 NMCSGRIELKFGRWGTVCDNFDNFDHASICRLECGSAVSFSGNSFGSGPIWF
 DDLICNGNESALWNCKHOGWKHNDHAEDAGVICSKGADLSRLVDGVTCSGRLEV

RFOEWGTCDDGDWSDAAVACKQLCQCPATVTAIGRVNASKGFGHILWLSVSCQGH
 PAYWCKHHEWGRHYCNHEDGAVTSDGSLRLRLRGGRSACAGTVEIQLLGV
 CDWGLKEADVVCRLQCGSALTKTQVYVSKIQTATWLFSSCNGNETSLMDCKNW
 QWGLTCDHYEAKITCSAHREPRLVGGDIPCSGRVEVKGHDITWGSICDSDFSLAAS
 VLCELOCTGVVSTILGGAHFGEGNQLWAEFOCEGHESHLCLCPVAPRPGTCSHSR
 DGVVCSRYTEIRLVNKGTPCEGRVELKTIGAWGSLCNSHWDIEDAHLVCOOLKCGVA
 LSTPGGARFGKNGQIWRHMFHCTGTQEHMGDCPVTALGASLCPSEQVAVSYCSGNS
 QTLSSNSSLGTPRTPIPEESAVACIESQGLRLVNGGRCAGREIYHGGSWGTICD
 DSWLSDAHVVCRLQCGEATNATGSAHFGEGTGTPIWLDKMKCNKESRWQCHSHGW
 GQONCRHKEDAGVTCSEFMSLRITSEAREACAGRLVFNWANGTGVKSKMSSTTVG
 VVRCLOCADKGINPASLDKAMSI PMWVNDVOCPKGPDLDLWQCPSPWKRLLASPS
 ETWITCDNKIRLOEGPTSCSGRVEIWHGSGKTVCDSDWLDLDAQVVCQOLCGCPALK
 AFRDEFQGTGTPIWLNKVEKNGNSSLMDPCPARWGHSECGHEDAAVNCITDLSVQK
 TPQKATGRSSRSSFTAVGILGVLLAIFALFLLFKRRQRORLAVSSRGENLHVQ
 IQYREMNSSLNADDLMLNSENSESADFAELISVSKFLPISGMEKEAILSTHEK
 ENGL"
 102..221
 222..3554
 /product="M130 antigen cytoplasmic variant 1"
 /evidence="experimental"
 polyA_signal 3760..3765
 polyA_signal 3764..3769
 polyA_signal 3768..3773
 polyA_site 3783
 BASE COUNT 1006 a 746 c 1074 g 960 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,77e-215 Length: 3786
 Score: 3441.50 Matches: 608
 Percent Similarity: 70.24% Conservatives: 166
 Best Local Similarity: 55.17% Mismatches: 311
 Query Match: 42.29% Indels: 17
 DB: 9 Gaps: 6
 US-09-759-130B-381 (1-1453) x HSM130AC1 (1-3786)
 QY 364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383
 DB 222 GGAACAGACAGAGAGCTGAGGTGTGTGTAACAAAGTAGCGGGAGAGTGGAA 281
 QY 384 ValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpIlyAsnGlnAla 403
 DB 282 GTGAAGTCCAGAGAGAGTGGGGAACGGTGTGTATTAATGGCTGGAGCATGCAAGCGTC 341
 QY 404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgAlaLys 423
 DB 342 TCTGTGATTGTAAACAGCTGGGATGTCCAACTGTATCAAAAGCCCTGGATGGCTAAT 401
 QY 424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443
 DB 402 TCCAGTGCAGGTTCCTGGAGCGCATTTGGATGATCATGTTCTTCGTGGGATGAGTCA 461
 QY 444 AlaLeuTrpAspCysThrTrpAspGlyLysAlaLysArgThr---CysPheArgArgSer 462
 DB 462 GCTCTTTGGATTGCAACATCATGATGGATGGGAAACATAGTACTACTACCAACAA 521
 QY 463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis 482
 DB 522 GATGCTGGAGTGCCTGCTCAGATGATGATCAATTTGGAATGAGGCTGAGCGGTGAGGG 581
 QY 483 SerProCysThrGlyArgLeuGluValLysTrpGlnGlyGluTrpGlyThrValCysHis 502
 DB 582 AATATGTGTCTGGAGAAATAGATCAATTTCCAAAGGACGGTGGGAGACAGTGTGTGAT 641
 QY 503 AspArgTrpSerThrArgAsnAlaAlaValCysLysGlnLeuGlyCysGlyLysPro 522
 DB 642 GATACTTCAACATAGATCATGCTGTCATTTGTAGACAACCTGATGATGAGTGTGCT 701
 QY 523 MetHisValPheGlyMetThrTrpPheLysGluAlaSerGlyProIleTrpLeuAspAsp 542
 DB 702 GTCAGTTTCTCTGTTTCATCTAATTTTGGGAAGGCTGTGGACCAATCTGTTGTTGAT 761


```
Db 2802 COATCATCTCCATGGAGAGAGACTGGCCAGCCCTCGGAGGACCTGGATCATCATGT 2861
Qy 1243 GluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrp 1262
Db 2862 GACACACAGATAAGACTTCAGAGAGACCCACTTCTGTTGGAGCTGTGGAGATCTGG 2921
Qy 1263 HisAlaGlySerTrpGlyThrValCysAspSerTrpAspLeuAlaGluVal 1282
Db 2922 CATGGAGTCTCTGGGGACAGTGTGATGACTCTTGAGACTTGGACGATCTCAGGTG 2981
Qy 1283 ValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAlaSerPheGly 1302
Db 2982 GTGTGTCAACAACCTGGCTGTGGTCCAGCTTTGAAGCATTCAGAGACGAGATTGTGT 3041
Qy 1303 GlnGlyThrGlyThrIleTrpLeuAspMetArgCysLysGlyAsnGluSerPheLeu 1322
Db 3042 CAGGGAGCTGGACCATATGGCTCAATGAAGTGAAGTCAAGAGGAATCAGTCTTCITG 3101
Qy 1323 TrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGly 1342
Db 3102 TGGGATTCTCTCGCAGAGCTGGGGCCATAGTAGTGTGGCACAAGAGACGCTGCA 3161
Qy 1343 ValArgCysSerGlyGlnSerLeuLys-----SerLeuAsnAlaSer 1356
Db 3162 GTGAATTCACAGATATTTCAGTCGAGAAACCCACAAAAGCCACACAGGTCGCTCA 3221
Qy 1357 SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPhe 1376
Db 3222 TCCCGTCAGTCATCTTTATTTCAGTCGGGATCCITGGGGTGTCTGTGGCCATTTC 3281
Qy 1377 Ile-----LeuPheIleuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArg 1394
Db 3282 GTCCGATTATTCTTCTGACTAAAAGCGAGACAGACAGACGCG-----CTTGCA 3332
Qy 1395 ValSerThrArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCys 1414
Db 3333 GTTCTCTCAAGGAGAGAACTTAGTCCACCAAAATCAATACCGGAGATGAATCTTGC 3392
Qy 1415 LeuLysArgGluAsp 1419
Db 3393 CTGAATGCAGATGAT 3407
RESULT 6
HSM130AC2 HSM130AC2 4950 bp mRNA: linear PRI 22-NOV-1996
LOCUS H.sapiens mRNA for M130 antigen cytoplasmic variant 2.
DEFINITION H.sapiens mRNA for M130 antigen cytoplasmic variant 2.
ACCESSION Z22970
VERSION Z22970.1 GI:312145
KEYWORDS antigen; antigen M130.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Micklem,K.K.
Direct Submission
Submitted (14-JUN-1993) Kingsley K.J. Micklem, Nuffield Department
of Pathology, University of Oxford, Level 1, Maternity Block, John
Radcliffe Hospital, Headington, Oxford, OX3 9DU, United Kingdom
2 (bases 1 to 4950)
Law,S.K., Micklem,K.J., Shaw,J.M., Zhang,X.P., Dong,Y., Willis,A.C.
and Mason,D.Y.
A new macrophage differentiation antigen which is a member of the
scavenger receptor superfamily
Eur. J. Immunol. 23 (9), 2320-2325 (1993)
JOURNAL 93380506
MEDLINE /db_xref="taxon:9606"
PUBMED /cell_type="Macrophage"
FEATURES
Location/Qualifiers
1..4950
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="Macrophage"
```

```
102..3572
/codon_start=1
/product="M130 antigen cytoplasmic variant 2"
/protein_id="CAA80543.1"
/db_xref="GI:312146"
/db_xref="SPTREMBL:Q07900"
/translation="MVLLDSGSDAFRRHFNLSPTTIVVLLLSACFVTSLSGGTDK
ELRLVDGKNCGRVEYKVQBEWGTVCNNGWSMEAVSVCNQLGCGPTAKPGWANS
AGSGRIWMDHVSRCRGNESALWDKHDGKHSNCTHOQDAGVTSDDGNLENRLTRGG
NMGSGRIEIKFQGRWGTVCDDNFNDHSAVICROLECGSAVSFCSSNFGSGPTWF
DDLICNGESALWNCKHOGWKHNCDAEDAGVICSKGADLSRLVDGVTGCGSGLV
RFQEWGTCDDGSDSDAAYACQQLGCTPTAVTAIGRYNASKGFGHILWDSVSCGHE
PAWCRHHEGKHFCNHNEDAVGTCSDGSDLELRGGRGRCAGTVEVEIQRLLGKV
QDGMGLKEADVVCVQLGSAKLTQYQVSKIQTATWLFSSCNGNETSLWDCKNW
QWGLTCDHYEEAKITCSAHREPLVGGDIPCSGRVEYHGDVTGSI CDSDFSLAAS
ILRELOQCTVVSILGGAHFGNGQIWAEEFCBGHESHSILCPVAPRPGCTCSHR
DVGVVCSRYTEIRLVNGTKPCGEGRVELTKLGWGLSCLNHWIDEDAHVLCQOLKCGVA
LSTGGARFGKNGQIWRHMFHCTGTQEHMGDCPVTAALGASLCPSEOVASVICSNGS
QTLSSCNSSSLGPTPTIPESAVACIESGQLRLVNGGRCAGRVEIHEGWSGTICD
DSDWLSDAHVVCRQLGCGEAINATGSAHFEGTGPIWDEMKCNKESRIWOCHSVG
QQCRKHEDAGVICSSEPMRLTSEAREACAGRLEVYNGAMGTGKSSNSETTVG
VVCQLGKADGKINPDSLKAMSTPMVNDVQPKPDTLWQCPSPSPWERKLAPSE
ETWITCDNKIRLOEPTSCSGRVEIWHGSGWGTVCDDSDWLDLDAOVVQQLCGPALK
AFKEATGCGTGPILNEVKCKGNESSLWDCPAREWGHSECGHKEADAQVNCNTDISVK
TPQKATGRSSRSQSFIAVGLIGVLLAIFVALFPLTKRQRQRQLAYSSRGLENLVHQ
IQRREMSNCLNADDDLLMNSGLWVGSIAGGPRSVAAVEAQTFFDFDKLKKKNVI
GSLDAYNGQE"
102..221
mat_peptide 222..3569
/product="M130 antigen cytoplasmic variant 2"
/revidence=experimental
polyA_signal 4924..4929
polyA_signal 4928..4933
polyA_signal 4932..4937
polyA_site 4947
BASE COUNT 1348 a 970 c 1277 g 1355 t
ORIGIN
Alignment Scores:
Pred. No.: 2,5e-214 Length: 4950
Score: 3429.00 Matches: 598
Percent Similarity: 71.27% Conservative: 161
Best Local Similarity: 56.15% Mismatches: 294
Query Match: 42.14% Indels: 12
DB: 9 Gaps: 4
US-09-759-130B-381 (1-1453) x HSM130AC2 (1-4950)
Qy 364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383
Db 222 GGAACAGACAGAGCTAGGCTAGTGGATGGTGAACAAAGTGTAGCGGAGAGTGGAA 281
Qy 384 ValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAla 403
Db 282 GTCAAGTCCAGAGAGTGGGGAACGGTGTCTAATATGCGTGGAGCATGGAACGGTCT 341
Qy 404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys 423
Db 342 TCTGTGATTGTGAACAGCTGGATGTCTCAACTGTATCAAAAGCCCTCGATGGGCTAAT 401
Qy 424 ProSerAsnGluAlaArgAspIleIleAsnSerIleSerCysThrGlyAsnGluSer 443
Db 402 TCCAGTCGAGGTCTCTGGAGCAGCTTTGGATGATCATGTTTCTTCTGCGGGAATGATCA 461
Qy 444 AlaLeuTyrAspCysThrTyrAspGlyLysAlaLysArgThr---CysPheArgArgSer 462
Db 462 GCTCTTTGGGATTGCAACATCATGATGGAGGGAACCATAGTACTGCTACTCACCACAA 521
Qy 463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis 482
Db 522 GATGCTGGAGTCACTGCTCAGATGGATCCCAATTTGGAATGAGGCTGACGCGTGGAGG 581
Qy 483 SerProCysTyrGlyArgLeuValLysTyrGlnGlyLeuTrpGlyThrValCysHis 502
```

Db 582 AATATGTCTCGAAGAATAGAGATCAAAATCCAAAGGACGGTGGGGAACAGTGTGTGAT 641
QY 503 AspArgTrpSerThrArgAsnAlaIaValValCysGlnLeuGlyCysGlyLysPro 522
Db 642 GATAACTTCAACATAGATGATGTCATTTGTAGACAACTGAATGTGGAAGTCT 701
QY 523 MethHisValPheGlyMetThrTyPheLysGluAlaSerGlyProIleTrpLeuAsp 542
Db 702 GTTCAGTTCTCTGTTCACTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTGATGAT 761
QY 543 ValSerCysIleGlyAsnGlnSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys 562
Db 762 CTTATATGCAACGAATAGTCACTGCTCTGGAACCTGCAACATCAAGGATGGGAAAG 821
QY 563 HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly 582
Db 822 CATAACTGTGATCATGCTGAGGATGCTGGAGTGTTCCTCAAGGAGGAGATCTGAGC 881
QY 583 LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyPheGlnGly 602
Db 882 CTGAGACTGCTAGATGGAGTCACTGAATGTTTCAGGAAGATTAGAGTGAGATTCGAAG 941
QY 603 ArgTrpGlyThrValCysAspGlyTrpAsnSerLysAlaIaValValCysSer 622
Db 942 GAATGGGGACAATATGCTGATGACGGCTGGGACAGTTACGATGCTGCTGGCATGCAAG 1001
QY 623 GlnLeuAspCysProSerIleGlyMetClyLeuGlyAsnAlaSerThrGlyTyr 642
Db 1002 CAACCTGGGATGTCCTCACTGCCGTCACAGCATGGTGGTCAAGCCCAAGGATTT 1061
QY 643 GlyLysIleTrpLeuAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys 662
Db 1062 GGACATCTGGCTTGACAGCGTTCTTCCAGGGACATGAACCTGCTGCTGGCAATGT 1121
QY 663 ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys 682
Db 1122 AAACACCATGAATGGGAAGCAATATTGCAATCAACAATGAAGATGCTGCGCTGACATGT 1181
QY 683 SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerArgCysAlaGlyLys 702
Db 1182 TCTGATGGATCAGATCTGAGCTTAAGACTTAGAGGTGGAGGACGCTGCTGGGACA 1241
QY 703 ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn 722
Db 1242 GTTGAGGTGGAGATTCAGAGACTGTTAGGAAGGTGTGTGACAGAGGCTGGGACATCAA 1301
QY 723 IleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGlu 742
Db 1302 GAAGCTGATGTGGTTGCAAGGACCTGGGATGGATGTGATCTGCACATCAAAACATCTTATCAA 1361
QY 743 ProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGly 762
Db 1362 GTGTACTCCAAATCCAGGCAACAAACATGCTGCTTTCTAAGTACTGCTAAGCGAAT 1421
QY 763 GluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsn 782
Db 1422 GAAACTTCTTTGGGACTGCAAGAACTGGCAATGGGTGGACTTACCTGTGATCACTAT 1481
QY 783 MetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMet 802
Db 1482 GAAGAAGCCAAAATTAACCTGCTCAGCCACAGGGAACCCAGACTGTTGGAGGGACATT 1541
QY 803 ProCysSerClyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSer 822
Db 1542 CCTGTGTCGGAGCTGTTGAAGTGAAGCATGTTGGACACGCTGGGCGCTCCATCTGTGATTCG 1601
QY 823 AspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIle 842
Db 1602 GACTTCTCTGGAACCTGCCACCGTTCATGTCAGGGAATTTACAGTGTGCACAGTTGTC 1661
QY 843 SerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPhe 862
Db 1661

Db 1662 TCTATCTCTGGGGAGCTCACTTTTGGAGAGGGAATGCACAGATCTGGGCTGAAGAAATTC 1721
QY 863 GlnCysGluCysSerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp 882
Db 1722 CAGTGTGAGGACATGATGCTCCATCTTCTCACTCTGCCAGTAGCACCCGCCCAAGAGGA 1781
QY 883 ThrCysIleHisSerArgGluValGlyValValCysSerArgTyThrAspValArgLeu 902
Db 1782 ACITGTAGCCACAGCAGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTG 1841
QY 903 ValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGly 922
Db 1842 GTGAATGGCAAGACCCGCTGTGAGGCGAGAGTGGAGCTCAAAACGCTTGTGCTCGGGA 1901
QY 923 SerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSer 942
Db 1902 TCCTCTGTAACTCTCACTGGACATAGAAGATCCCATGTTCTTCCACAGACCTTAAA 1961
QY 943 CysGlyThrAlaLeuSerThrThrGlyGlyLysTyIleGlyGluArgSerValArgVal 962
Db 1962 TGTGGAGTTGCCCTTTCTACCCAGGAGGAGCAGCTTTTGGAAAAGGAATGGTCAGATC 2021
QY 963 TrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThr 982
Db 2022 TGGAGGCATATGTTTCACTGACCTGGGACTGAGCAGCAGCATGGGAGATTTGCTCTGAAT 2081
QY 983 ValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGlySer 1002
Db 2082 GCTCTGAGTCTTCATTTATGTTCTTTCAGAGCAAGTGGCTCTGTAAATCTGCTCAGGAAC 2141
QY 1003 LeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyThrLeuSerAlaVal 1022
Db 2142 CAGTCCCAACACTGTCTCTGTCGCAATTCATCGTCTTTGGGCCCAACAAGCCCTACCATT 2201
QY 1023 ProGluCysSerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAsp 1042
Db 2202 CCAGAAGAAAGTGTGTGGCTGTGATAGAGAGTGTCAACTTCGCCCTGGTAAATGGAGGA 2261
QY 1043 SerArgCysAlaGlyArgValGluIleTyHisAspGlyPheTrpGlyThrIleCysAsp 1062
Db 2262 GTTCGCTGCTCTGGAGAGTAGAGATCTATCATGAGGGCTCTCTGGGCAACCATCTGTGAT 2321
QY 1063 AspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAla 1082
Db 2322 GACAGCTGGGACCTGAGTGTGATGCCACCTGTTTGGACACAGCTGGCTGTGGAGAGGCC 2381
QY 1083 PheAsnAlaThrValSerAlaHisPheGlyCysGlySerGlyProIleTrpLeuAspAsp 1102
Db 2382 ATTAATGCCACTGTTCTGCTCATTTTGGGAAGGAACAGGGCCCATCTGGCTGGATGAG 2441
QY 1103 LeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGln 1122
Db 2442 ATGAATGCAATGGAAAAGAAATCCCGCATTTGGCAGTGGCCATTCACAGCGTGGGGCAG 2501
QY 1123 HisAsnCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArg 1142
Db 2502 CAAAATGGCAGGCACAAAGGAGGATGCGGGAGTTATCTGCTCAGAATTCATGCTCTGAGA 2561
QY 1143 LeuTyrsSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyAsnGly 1162
Db 2562 CTGACCACTGAAGCCACAGCAGAGAGGCGCTGTCAGGGGCTCTGGAAGTTTTTTTACAATGGA 2621
QY 1163 ThrTrpGlySerValGlyArgArgAsnIleThrAlaIleAlaGlyIleValCysArg 1182
Db 2622 GCTTGGGGCACTGTTGGCAAGAGTAGCATGCTGTAACCCACTGTTGGGTGGTGTGGCAGG 2681
QY 1183 GlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySer 1202
Db 2682 CAGCTGGCTGTGCAGACAAAGGGAATAATCAACCTGCATCTTTAGACAGGCCATGTCC 2741
QY 1203 GlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCys 1222
Db 2742 ATTCCTCATGTGGTGGCAATGTTTCAGTGTCCAAAAGGACCTGCACAGCTGTGGCAGTGC 2801

```

Qy 1223 LeuSerAlaProTTPGluArgArgIleSerSerProAlaGluGluThrTrpIleThrCys 1242
Db 2802 CCATCATCTCCATGGAGAGAGACTGCCAGCCCTCGGAGGAGACTGGATCAGCATGT 2861
Qy 1243 GluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrp 1262
Db 2862 GACACACAGATAAGACTTCAGAGAGCCCACTTCCTGTTCTGGACGTGGGATCTGG 2921
Qy 1263 HisAlaGlySerTrpGlyThrValCysAspSerTrpAspLeuAlaGluAlaGluVal 1282
Db 2922 CATGGAGTTCTCTGGGGACAGTGTGATCAGCTCTGGGAGCTGGAGGATCTCAGGTG 2981
Qy 1293 ValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAlaSerPheGly 1302
Db 2982 GTGTGTCACCAACTGGCTGTGGTCCAGCTTTGAAAGCATTCACAAAGACAGAGTTTGGT 3041
Qy 1303 GlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeu 1322
Db 3042 CAGGGAGCTGACCGCATATGCTCAATGAAGTGAAGTCAAGGGAATGAGTCTTCCTTG 3101
Qy 1323 TrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGly 1342
Db 3102 TGGGATTCTCTGCCAGAGCTGGGGCCATAGTAGTGTGGGCACAGAGAGAGCGCTCA 3161
Qy 1343 ValArgCysSerGlyGlnSerLeuLys-----SerLeuAsnAlaSer 1356
Db 3162 GTGAATTCACACAGATATTTTCAGTGCAGAAAAACCCACAAAAAGCCACACAGGTCGCTCA 3221
Qy 1357 SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPhe 1376
Db 3222 TCCCGTCAGTCATCTTTATTCAGTGGGATCTGGGATCTTGGGTTGTCTCTTGGCCATTTTC 3281
Qy 1377 Ile-----LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArg 1394
Db 3282 GTCGATTATTTCTTCTGACTTAAAGCGAGACAGACAGACAGCGG-----CTTGCA 3332
Qy 1395 ValSerThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCys 1414
Db 3333 GTTCTCTCAAGAGGAGAGAACTTAGTCCACCAAAATTCATACCGGGAGATGAATTCCTTCG 3392
Qy 1415 LeuLysArgGluAsp 1419
Db 3393 CTGAATGCAGATGAT 3407
RESULT 7
HSM130AE HSM130AE 3802 bp mRNA linear PRI 22-NOV-1996
DEFINITION H.sapiens mRNA for M130 antigen extracellular variant.
ACCESSION Z22971
VERSION Z22971.1 GI:312147
KEYWORDS antigen; antigen M130.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Direct Submission
JOURNAL Submitted (14-JUN-1993) Kingsley K.J., Micklem, Nuffield Department
of Pathology, University of Oxford, Level 1, Maternity Block, John
Radcliffe Hospital, Headington, Oxford, OX3 9DU, United Kingdom
REFERENCE
AUTHORS Law, S.K., Micklem, K.J., Shaw, J.M., Zhang, X.P., Dong, Y., Willis, A.C.
and Mason, D.Y.
TITLE A new macrophage differentiation antigen which is a member of the
scavenger receptor superfamily
JOURNAL Eur. J. Immunol. 23 (9), 2320-2325 (1993)
MEDLINE 93380506
PUBMED 8370408
FEATURES
Location/Qualifiers
source
1..3802
/organism="Homo sapiens"

```

CDS

```

/db_xref="taxon:9606"
/cell_type="Macrophage"
102..3551
/codon_start=1
/product="M130 antigen extracellular variant"
/protein_id="CA80544.1"
/db_xref="GI:312148"
/db_xref="SPTREMBL:Q07901"
/translaton="MVLDESGSADFRRHFVNLSPTFTVVLLLSACFVTSLSGGTOK
ELRLVDENKSCSRVEYKQVEWGTVCNNGSMSEAVSICNOLGCPATIAKAPGWANS
AGSGRIWMDHVSCRGNESALWDCKHDGKHSNCTHOQDAGVTCSDGNSNLEMRITRG
NMCSGRIEIKFOGRWGTVCDDNFENIHASVICROLECGSAVSFGSGNSFEGSGPIWF
DDLICNGESALWNCKHQKNCDAEDAGVTCCKGADLSRLVDGVTECSGRLEV
RFGWGTICDDGWDSTDAAVACKQLGCPATVTAIGRVNASKGFGHILWLSVSCQGH
PAYQCKHHEWGHYCNHEDAGVTCSDGSDLELRLLRGSGRCAGTVEVEIURLGKV
CWGLTKLEADVVCROQLGCSALKTSQVTSKIATNTWLFSSCNGNETSLWDCNN
QWGLTGDHYEAKITCSAHREPLRGVDIPCSGRVEVKHGTWGSICDSDFLEAAS
VLCRELQGTVVSTLGAHREGGQIWAEEFCQEGHESHLSCPCVAPRPGTCSHSR
DVGWVSSKTQKTSLGSYTVKGTGLSHSCLFLKPCLLPGYTEIRLVNCKTPCEGRV
ELKTLGWSGLCSNHWIDIEAHVLCQQLKCGVALSTPGGARFGKNGQIWRHMFCTG
TEHMGDGPVTLGALGSCPSEQVAVSVCISGNOSQTLSSCNSSSLGTPRTPIEESAVA
CIBSGQLRVNGGRCAGRVEIYHEGWSGTICDDSDLSDAHVYCRQLGCEALNATG
SAHFEGAGTPIWLDENKNGKESRWOCHSHGWGQONCRKEDAGKINPASTDKRAMSI
PMWVDVQCPKGPDTLMQCPSSPWEKRLASPSSEETWITCDNKIRLOEGPTSCSRVEI
WHGSGWTVCDDSDLDLDAQVVCQQLGCGPALKAFKAEFGQGTGPWLNEVVKNGVE
SSLWDCPARRHGHSECKEDAAVNCQDTSVQKTPQKATGRSRSQSFIAVLGLVY
LLAIFVALFELTKRRQRQLAVSSRGENLVHQIYQREMNSCLNADLDLMSNGGHS
EPH"
sig_peptide 102..221
mat_peptide 222..3548
/product="M130 antigen extracellular variant"
/evidence=experimental
polyA_signal 3776..3781
polyA_signal 3780..3785
polyA_signal 3784..3789
polyA_site 3799
BASE COUNT 1009 a 751 c 1081 g 961 t
ORIGIN
Alignment Scores:
Pred No.: 2,41e-212 Length: 3802
Score: 3396.50 Matches: 597
Percent Similarity: 69.03% Conservative: 161
Best Local Similarity: 54.37% Mismatches: 295
Query Match: 41.74% Indels: 45
DB: 9 Gaps: 5
US-09-759-130B-381 (1-1453) x HSM130AE (1-3802)
Qy 364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383
Db 222 GGAACAGACAAGGAGCTGAGGCTAGTGGATGCTGTAACAAAGAGTAGCGGAGAGTGGAA 281
Qy 384 ValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAla 403
Db 282 GTGAAGTCCAGGAGGAGTGGGACGCTGTATAATATGGCTGGAGCATGGAACGGCTC 341
Qy 404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys 423
Db 342 TCTGTGATTTGTAAACACCTGGATGCTCAACTGCTATCAAGAGCCCTGGATGGCTAAT 401
Qy 424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443
Db 402 TCCAGTGCAGGTTCTGGACGCTATTTGGATGATCATGTTCTTCTGCTGGGAATGAGTCA 461
Qy 444 AlaLeuTrpAspCysThrTyrrAspGlyLysAlaLysAlaArgThr---CysPheArgArgSer 462
Db 462 GCTCTTTGGGATTCGAAACATGATGGATGGGAAAGCATGACTACTACTACCAACAA 521
Qy 463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuValGlyAlaHis 482
Db 522 GATGCTGAGTGACCTGCTCAGATGGATCCCAATTTGGAAATGAGGCTGACCGCTGGAGG 581

```


Qy	843	SerLeuSerValcylAspHis	shgLyLysGlyAsnGlyLeuThrTrpAlaGluLysphe	862
Db	1662	TCTATCTCGGGGGAGCTC	ACTTTGGAGAGGAAATGGACAGATCTGGGCTGAAGAATTC	1721
Qy	863	GlnCysGluGlySerGluThrHis	LeuAlaLeuCysProIleValGlnHisProGluAsp	882
Db	1722	CAGTGTGAGGGACATGAGTCC	CATCTTTCACCTGCCCAGTAGCACCCCGCCCAAGAGGA	1781
Qy	883	ThrCysIleHisSerArgGlu	ValGlyValValCysSerArg-	896
Db	1782	ACTTGTGACCCACAGCAGGATGT	GTGAGTAGTCTCTCAAGTAAGACCCAGAAAAACATCT	1841
Qy	896	-----	-----	896
Db	1842	TTAATTGGTCTTATACT	TGTGAAGAGGACAGGGTTTAGGAGGTCATPAGTGTCTTTTCTA	1901
Qy	897	-----	-----TyrThrAspValArgLeuValAsnGlyLysSerGlnCys	909
Db	1902	AGCCCTGTCTCCTTCC	AGNATACACAGAAATTCCTCTGGTGAATGGCAGACCCCGTGT	1961
Qy	910	AspGlyGlnValGluIleAsn	ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrp	929
Db	1962	GAGGCGAGAGTGGAGCTCAAA	CGCTTGTGCTGGGATCCCTCTGTAACTCTCACTCTCAGTCG	2021
Qy	930	AspProGluAspAlaArgVal	LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThr	949
Db	2022	GACATAGAAGATGCCCATGTCT	TTTGGCCAGCAGCTTAATGTGGAGTTGGCCTTTCTACC	2081
Qy	950	ThrGlyGlyLysTrpIleGly	GluArgSerValArgValTrpGlyHisArgPheHisCys	969
Db	2082	CCAGGAGGACAGCTTTTGG	AAAGAAATGTCAGATCTGGAGGCATATGTTTCACATGC	2141
Qy	970	LeuGlyAsnGluSerLeuLeu	AspAsnCysGlnMetThrValLeuGlyAlaProCys	989
Db	2142	ACTGGGACTGAGCAGCACAT	GGAGATGCTCTGTAACCTGCTAGGTGCTTCATTATGT	2201
Qy	990	IleHisGlyAsnThrValSer	ValIleCysThrGlySerLeuThrGlnProLeuPhePro	1009
Db	2202	CCCTCAGCAGCAAGTGCCCT	CTGTAAATCTCTCAGAAACAGCTCCCAACACTGTCGTCG	2261
Qy	1010	CysLeuAlaAsnValSerAsp	ProTyrLeuSerAlaValProGluGlySerAlaLeuIle	1029
Db	2262	TGCATTCATCGTCTTTGG	CCCCCACAAAGCCCTACCATTCACAGAGAAAGTGCTGTGCC	2321
Qy	1030	CysLeuGluAspLysArgLeu	ArgLeuValAspGlyAspSerArgCysAlaGlyArgVal	1049
Db	2322	TGCATAGAGAGTGGTCAACT	TTCGCCCTGGTAAATGGAGGAGTGCCTGCTGGGAGAGTA	2381
Qy	1050	GluIleTyrHisAspGlyPhe	ThrTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAsp	1069
Db	2382	GAGATCTATCATGAGGCTCT	GGGCGCCATCTGTGTAGACAGCTGGGACCTCAGTGTAT	2441
Qy	1070	AlaHisValValCysGlnLys	LeuGlyCysGlyValAlaIlePheAsnAlaThrValSerAla	1089
Db	2442	GCCACGTGTGTTCAGACAG	CTGGGCTGTGGAGAGGCCATTAATGCCACTGGTTCGTCT	2501
Qy	1090	HisPheGlyGluGlySerGly	ProIleTrpLeuAspAspLeuAsnCysThrGlyThrGlu	1109
Db	2502	CATTTTGGGAAGAAACAG	GGCCCCATCTCGCTGGATGAGATGAATGCAATGGAAAGAA	2561
Qy	1110	SerHisLeuTrpGlnCysPro	SerArgGlyTrpGlyGlnHisAspCysArgHisLysGlu	1129
Db	2562	TCCCGCATTTTGGCAGTGC	CATTCACACGCGCTGGGGCGAGCAAAATTCGACGCACAAAGG	2621
Qy	1130	AspAlaGlyValIleCysSer	GluPheThrAlaLeuArgLeuTyrSerGluThrGluThr	1149
Db	2622	GATCGGGAGATTATCTGCT	CAGAAATTCATGCTCTGTAGACTGACCCAGTGAAGCCAGCAGA	2681
Qy	1150	GluSerCysAlaGlyArgLeu	GluValPheTyrAsnGlyThrTrpGlySerValGlyArg	1169
Db	2682	GAGCCCTGTCCAGGGCGT	CTGGAGCTTTTACATGTGGAGCTTGGGGCAGCTGTGGCAG	2741
Qy	1170	ArgAsnIleThrThrAlaIle	AlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsn	1189

```
Db 2742 AGTAGCATGCTCTGAACCAACTGTGGGTGTGGTGTGACGAGCTGGGTGTGACAGAAA 2801
Qy 1190 GlyValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAsp 1209
Db 2802 GGGAAATCAACCCCTGATCTTTAGACAGGCCATGTCATCCATCCATGGGTGGACAA 2861
Qy 1210 IleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArg 1229
Db 2862 GTTCAGTGTCCAAAGGACCTGACACGCTGTGGCAGTGGCCATCATCTCCATGGGAGA 2921
Qy 1230 ArgIleSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArgValArg 1249
Db 2922 AGACTGGCCAGCCCTCGGAGAGACCTGGATCAGATGATGACACAAAGATAGACTTCAG 2981
Qy 1250 GlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThr 1269
Db 2982 GAAGAGCCCATCTCTGTCTGGAGCTGGAGATCTGCGATGGAGGTTCCTGGGGGACA 3041
Qy 1270 ValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCys 1289
Db 3042 GNGTGTGATGACTCTTGGGACTTGGACATGCTCAGGTGGTGTGTCACAACTTGGCTGT 3101
Qy 1290 GlySerAlaLeuAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrp 1309
Db 3102 GTCAGCATGTTTGAAGCATTCACAAAGACAGAGTTGTGTCAGGGGACTGGACCGATATGG 3161
Qy 1310 LeuAspAspMetArgCysLysGlyAsnGlnLeuTrpPheLeuTrpAspCysHisAlaLysPro 1329
Db 3162 CTCATGAGTGAAGTGAAGGAAAGTGAAGTCTCTCTGTGGGATGTCCTGCCAGAGCC 3221
Qy 1330 TrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSer 1349
Db 3222 TGGGGCCATAGTGTGGGACAAAGGAGAGCTGCGTCAATGTCACAGATATTTCA 3281
Qy 1350 LeuLys-----SerLeuAsnAlaSerSerGlyHisLeuAlaLeuIle 1363
Db 3282 GTGCAGAAACCCCAAAACCCCAACAGGCTCATCCGTCAGTCATCCTTTAT 3341
Qy 1364 LeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPheIle-----LeuPheLeuThr 1381
Db 3342 GCAGTCGGGATCCTTGGGGTGTCTGTGGCCATTTTCGTCGCATATTCTCTTGACT 3401
Qy 1382 TrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgGly 1401
Db 3402 AAAAAGCGAAGACAGACAGACGG-----CTTGCAAGTTTCTCCTCAAGAGAGAGA 3452
Qy 1402 SerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAsp 1419
Db 3453 TTAGTCCACCAATTCATACCGGGAGATGAATCTTGCTGATGACAGATGAT 3506

RESULT 8
LOCUS AX281768 3707 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 177 from Patent WO0177389.
ACCESSION AX281768
VERSION AX281768.1 GI:16609019
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
REFERENCE
AUTHORS Shiffman,D., Somogyi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,
Mikita,T. and Tai,J.
TITLE Genes expressed in foam cell differentiation
JOURNAL Patent: WO 0177389-A 177 18-OCT-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1..3707
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 236574.12"
```

```
BASE COUNT 984 a 731 c 1062 g 930 t
ORIGIN
Alignment Scores:
Pred. No.: 4.38e-211 Length: 3707
Score: 3377.00 Matches: 598
Percent Similarity: 71.00% Conservatives: 161
Best Local Similarity: 55.94% Mismatches: 294
Query Match: 41.50% Indels: 16
DB: 6 Gaps: 4

US-09-759-130B-381 (1-1453) x AX281768 (1-3707)
Qy 364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSer-GlyArgValG1 383
Db 222 GGAACAGACAAAGGAGCTGAGGCTAGTGTGCTGAAACAACTAGCGGGAGAGTGA 281
Qy 383 uValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAl 403
Db 282 AGTGAAGTCCAGGAGGAGTGGGAAACGCTGTGTAATAATGGCTGGACATCGAAGCGGT 341
Qy 403 aLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgAlaLys 423
Db 342 CTCCTGTGATTTGTAACAGCTGGGATGTCCTCAACTCTATCAAAAGCCCTGGATGGCTAA 401
Qy 423 sProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSe 443
Db 402 TTCAGTGCAGGCTTTCGACGACATTTGGATGGATCATGTTCTTGTGCTGGGAATGAGTC 461
Qy 443 rAlaLeuTrpAspCysThrTyrrAspGlyLysAlaLysArgThr---CysPheArgArgSe 462
Db 462 AGCTCTTTGGGATTCACAAACATGATGGATGGGAAAGACATAGTAATCTACTACCAACA 521
Qy 462 rAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHi 482
Db 522 AGATCTGGAGTGACCTGCTCAGATGGATCCAAATTTGGAAATGAGCGTCACCGTGGAGG 581
Qy 482 sSerProCysTyrrGlyArgLeuGluValLysTyrrGlnGlyGluTrpGlyThrValCysHi 502
Db 582 GAATATGCTTCTCGAAGAATAGAGATCAAAATTCGAAGGACGCTGGGGAACAGTGTGTA 641
Qy 502 sAspArgTrpSerThrArgAsnAlaValValCysLysGlnLeuGlyCysGlyLysPr 522
Db 642 TGATAACTTCAACATAGATCATGCTCATTTGTAGACAACTGAATGATGGAGAGTGC 701
Qy 522 oMetHisValPheGlyMetTyrrPheLysGluAlaSerGlyProIleTrpLeuAspAs 542
Db 702 TGTCAGTTTCTCTGCTCATCTAATTTTGGAGAAGGCTCTGACCAATCTGCTTTCATCA 761
Qy 542 pValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLy 562
Db 762 TCTTATATCCAAACGGAAATGATGACGCTCTCTGGAACTGCAACATCAAGAGTGGGAAA 821
Qy 562 shiAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpG1 582
Db 822 GCATAACTGTGATCATGCTGAGGATGCTGGAGTGTGCTCAAGGGAGCAGATCTGAG 881
Qy 582 yLeuArgLeuValGlyLysSerAsnArgCysSerGlyArgLeuGluValTyrrPheGlnG1 602
Db 882 CCGTAGACTGGTGTAGATGGAGTCACTGAATGTTTCAGAGAGATTAGAAGTGGAGTCCAAG 941
Qy 602 yArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValValCysSe 622
Db 942 AGAATGGGGGACAAATGTGATGACGCTGGGACAGTTACGATGCTGTGCTGGCATGCA 1001
Qy 622 rGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTy 642
Db 1002 GCAACTGGATGTCCAACCTGCGCTCACAGCCATTTGGTCAGTTAAGCCAGTAAGGAT 1061
Qy 642 rGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCy 662
Db 1062 TGGACACATCTGGCTTGACAGCGTTTCTTGCCAGGAGACATGAACCTGCTGTCTGCGCAATG 1121
```


Qy 662 sArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys 682
Db 1122 TAAACACCAGTAAATGGGAAAGCATTTATGCAATCAATCAATGAGTCTGGCGTGACATG 1181
Qy 682 sSerAspAlaSerAspMetGluLeuArgLeuValGlyCysSerArgCysAlaGlyLy 702
Db 1182 TTCTGATGGATCAGATCTGGAGCTTAAGACTTAGAGGTTGGAGCGCTGTGTGGGAC 1241
Qy 702 sValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAs 722
Db 1242 AGTTGAGCTGGAGATTCAGAGACTGTAGGAAGGTGTGTGACAGAGCTGGGGAGTAA 1301
Qy 722 nIleAlaGluValValCysArgGlnLeuGlyCysGlySerAlaIleArgValSerArg 742
Db 1302 AGAAGCTGATGTGTTTGGAGGAGCTGGGATGTGGATCTGCACCTCAAAACATCTTATCA 1361
Qy 742 uProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyG 762
Db 1362 AGTGACTCCAAAATCCAGGCCAACAAACACATGGCTGTTTCTAAGTAGCTGTAAACGAAA 1421
Qy 762 yGluAlaSerLeuTrpAspCysIleArgTTPGluTrpLysGlnThrAlaCysHisLeuAs 782
Db 1422 TGAACACTCTCTTGGGACGTGCAAGACTGGCAATGGGGTGGACTTACCTGTGATCACTA 1481
Qy 782 nMetGluAlaSerLeuLeuCysSer-AlaHisArg-GlnProArgLeuValGlyAlaAsp 801
Db 1482 TGAAGAAGCCAAATACCTGCTCAGGCCACACAGGGGACCCAGACTGGTTGGAGGGAC 1541
Qy 802 MetProCysSerGlyArgValGluValValHisAlaAspThrTrpArgSerValCysAsp 821
Db 1542 ATTCCTCTGTTGGACGTGTGAAGTGAAGCATGGTGACAGCTGGGGCTCCATCTGTGAT 1601
Qy 822 SerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAla 841
Db 1602 TCACACTCTCTCTGGAGCTGCCAGCTTATGCAAGGAAATACAGTGTGGCACAGTT 1661
Qy 842 IleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlyLys 861
Db 1662 GTCTCTATCTCTGGGGGAGCTCACTTTGGAGAGGGAAATGACAGACTCTGGGCTGAAGAA 1721
Qy 862 PheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGlu 881
Db 1722 TTCAGTGTGAGGAGCATGATGCCATCTTCACTCTGCCAGTAGCACCCGCCCCAGAA 1781
Qy 882 AspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrTrpAspValArg 901
Db 1782 GGAATTTGTACCCACACAGCGGATGTGGAGTAGTCTGCTCAAGATACACAGAAATTCG 1841
Qy 902 LeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrp 921
Db 1842 TTGTGAATGGCAAGACCCCGTGTGAGGCGAGAGTGGAGCTCAAAACGCTTGGTGCCTGG 1901
Qy 922 GlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeu 941
Db 1902 GGATCCCTCTGTAACTCTCACTGGGACATAGAAGATGCCATGCTTCTTGGCAGCAGCTT 1961
Qy 942 SerCysGlyThrAlaLeuSerThrGlyCysLysThrIleGlyGlyValArgSerValArg 961
Db 1962 AAATGTGGAGTGGCCCTTCTTACCCAGGAGGACAGCTTTTGGAAAGAAATGGTCAG 2021
Qy 962 ValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMet 981
Db 2022 ATCTGGAGGATATGTTTCACTGACCTGGGACTGAGCAGCAGCATGGGAGATTGTCCTGTA 2081
Qy 982 ThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGly 1001
Db 2082 ACTGCTCTAGGTGCTTCAATATGCTCTACAGCAAGTGGGCTCTGTAATCTGTCTCAGGA 2141
Qy 1002 SerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAla 1021
Db 2142 AACAGTCCCAACACATGTCCTGTGCAATTCATCTGCTTTGGGCCCAACAGCCCTACC 2201
Qy 1022 ValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGly 1041

Db 2202 ATTCACGAAGAAGTGTGTGGCTGTACAGAGAGTGGTCAACTTCGCTGGTAAATGGA 2261
Qy 1042 AspSerArgCysAlaGlyArgValGluIleThrHisAspGlyPheTrpGlyThrIleCys 1061
Db 2262 GGAGTCTGCTGTGTGGGAGATAGATCTATCATGAGGCTCTGGGGACCATCTGT 2321
Qy 1062 AspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyVal 1081
Db 2322 GATGACAGCTGGGACCTGAGTATGCCACGTGGTGTGACAGACAGCTGGGCTGTGGAG 2381
Qy 1082 AlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAsp 1101
Db 2382 GCCATTAAATGCCACTGTGCTTCTGCTCATTTGGGAAGAACAGGCCCATCTGGCTGAT 2441
Qy 1102 AspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGly 1121
Db 2442 GAGATGAATGAATGAAGAAATCCGCAATTTGGCAGTGGCATCACAGCTGGGG 2501
Qy 1122 GlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeu 1141
Db 2502 CAGCAAAATTCAGGCACACAGGAGATGCCGGAGTTATCTGCTCAGAAATTCATGCTCTG 2561
Qy 1142 ArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsn 1161
Db 2562 AGACTGACCACTGAAGCCAGCAGAGAGGCGCTGTGCAGGCGCTCTGGAAGTTTTACAA 2621
Qy 1162 GlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCys 1181
Db 2622 GGAGCTTGGGCACTGTGGCAAGAGTAGTCTGTAACACCACTGTGGGTGGTGTGGC 2681
Qy 1182 ArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGly 1201
Db 2682 AGGCAGCTGGCTGTGCAGACAAAGGGAATCAACCTGTCATCTTTAGACAGGCCATG 2741
Qy 1202 SerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrp-G 1221
Db 2742 TCCATTCCCATGTGGTGGCAATGTTCACTGCTCCAAAGGACCTGACACGCTGTGGCA 2801
Qy 1221 nCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleTh 1241
Db 2802 GTGCCATCATCTCCATGGGAGAGAGACTGCCAGCCCTCGGAGGAGACTGGATCAC 2861
Qy 1241 rCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluI 1261
Db 2862 ATGTGACAAACAAGATAAGACTTCAGGAAGGACCCACCTTCCTGTTCTGGACGTTGGAG 2921
Qy 1261 eTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaG 1281
Db 2922 CTGCATGGAGTTCCTGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCA 2981
Qy 1281 uValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPh 1301
Db 2982 GGTGGTGTCTCAACAACCTTGGCTGTGGTCCAGCTTTGAAGACATTCAAAGAGCAGAGT 3041
Qy 1301 eGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPh 1321
Db 3042 TGGTCAAGGAGTGGACCGATATGGCTCAATGAAGTGAAGTGAAGGAAGATGAGTCTTC 3101
Qy 1321 eLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAl 1341
Db 3102 CTTCGTGGATGTCCTGCCAGACGCTGGGGCCATAGTGTGGGACACAAAGGAAGACGC 3161
Qy 1341 aGlyValArgCysSerGlyGlnSerLeuLys-----SerLeuAsnAl 1355
Db 3162 TGCAGTGAATTCACAGATATTTTCAGTGCAGAAAAACCCCAAAAGAGCAGAGTTCG 3221
Qy 1355 aSerSerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLe 1375
Db 3222 CTATCCGCTCAGTCATCTTATTTGAGTGGGATCTTGGGGTGTCTGTGGCCAT 3281
Qy 1375 uPheIle-----LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLe 1393
Db 1393

```

Db 3282 TTTCGTCCGATTATCTCTTCTGACATAAAAGCGAAGACAGACAGACAGCGG-----CT 3332
QY 1393 uArgValSerThrArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluTh 1413
Db 3333 TGCAGTTTCTCCTCAAGAGAGAGAACTTAGTCCACCAAAATTCATACCGGAGATGAATC 3392
QY 1413 rCysLeuLysArgGluAsp 1419
Db 3393 TTGCCTGAATGCAGATGAT 3411

RESULT 9
LOCUS AF274883 4405 bp mRNA linear ROD 10-MAY-2001
DEFINITION Mus musculus macrophage hemoglobin scavenger receptor CD163
ACCESSION AF274883
VERSION AF274883.1 GI:13236130
KEYWORDS
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 4405)
AUTHORS Schaefer, D.J., Boret, F.S., Hongegger, A., Poehler, D., Linscheid, P.,
Staege, H., Muller, C., Schoedon, G. and Schaffner, A.
TITLE Molecular cloning and characterization of the mouse CD163
homologue, a highly glucocorticoid-inducible member of the
scavenger receptor cysteine-rich family
JOURNAL Immunogenetics 53 (2), 170-177 (2001)
MEDLINE 21243729
PUBMED 11345593
REFERENCE 2 (bases 1 to 4405)
AUTHORS Schaefer, D.J., Boret, F.S. and Schaffner, A.
TITLE Direct Submission
JOURNAL Submitted (03-JUN-2000) Departement of Medicine, Medizinische
Klinik B, University Hospital Zurich, Ramistrasse 100, Zurich, ZH
8091, Switzerland
FEATURES
location/Qualifiers
1..4405
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="6"
/map="15.43 centirays distal from the D6Mit300"
85..3430
/feature="FAT; member of the scavenger receptor cysteine rich
type B superfamily; transmembrane protein type I"
/codon_start=1
/product="macrophage hemoglobin scavenger receptor CD163
precursor"
/protein_id="AAK16065.1"
/db_xref="GI:13236131"
/translation="MGHHRVLLGGAGSGPGKRFVHLGFFVAVVSSLLSASAVTNAPG
EMKRLRAGGENNCSGRVELEKIDKWTGVCNGSMNVEVSVQQLGCPSTIKALGW
ANSSAGSYIWMDCVCTGNSALWDCKHGDGKHNCTHEKADGVTQCGSLNLEMLV
NSAGHCLGRVEIKFQGWGTVCDNDFSKDHASVTCOLGCGSALSFSGESAKLGAGC
PIWDDLACNENESALWDCKHGRGKHNCDHEDVGVICLEGADLSRLVDCVRCSG
RLEVFQGEWGTVDNDLDRDASVYQVLCGCTAISAIRVNASSEGGQIWLNDISC
EGHEATLNECHQWEMGRYCHREDDVQVCTSDGADLELRVLGGSGRCAIVEVEIKL
TKMGLDNLADADVCRQLCGSALQTAQIYKSTGATNTWLPFGSCNGNETTFWQ
CKMOWGLSCDNFEAKVCTSGHREPRLVGGEIPCSGRVEVKHGVMSYCDFDLST
EASVYVRELQCGTVSVILGGAHFGESGQWGEFQCGSGDESHLSLCSVAPPLDRTC
THRSVYVCYRIIDIRLAGVSSCEGRVELKTLGANGPLCSHMDMDAHVLCQKL
CGVAQSIPEGAFHGAGGVWSHMFCTGTEEHIGDCLMTALGAPTCSEGGVAVQIC
GNQSTLLPCSLSPVOTSTIPKESEVPCIASGQLRLVGGRCAGRVEVHEGWS
GTVCDDNDMDTDANVYQQLDGVAINATGSAIFEGAGAIWLDVICTGKSHIWOQ
HSHGWRNCRKHEKADGVICFEFMSRLTNEAKHENTCTGRLEVFYNGTWGSGISNMS
PTTVGVCRLQGLCADNVTKPTPSDKTPSRMVDVRCQPKGVDTLMOCPSSPMKQRO
ASPSQSEWIIIDKNKIRLQEGHTDCSGRVEIWHKGSWGTVCDDSDWLDNDKAVKQGLG
CQAVKALRAEAFPGTPIWLNEIKRGNESLWDCPKAPKWSHSDCKEDASIQCL
VOYQEMDSKADLDLLKSSGVIQRTHEKENDL"
85..198
sig_peptide
misc_feature 3220..3288

```

```

BASE COUNT 1174 a 884 c 1246 g 1101 t
ORIGIN
Alignment Scores: 3.85e-204 Length: 4405
Pred. No.: 3272.00 Matches: 582
Score: 68.11% Conservative: 174
Percent Similarity: 52.43% Mismatches: 332
Best Local Similarity: 40.21% Indels: 22
Query Match: 10 Gaps: 9
DB: 9

US-09-759-130B-381 (1-1453) x AF274883 (1-4405)
QY 320 GlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrp 339
Db 91 GGACACAGAAATGGTCTTCTTGGAGGTGCTGCATCTCCTGGTGTGTAAGAGTTGCTG--- 147
QY 340 AspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerVal 359
Db 148 -----CATCTAGGTCTTCTTCTTGGCTGTGAGCTCACTTCTCAGTCCCTCTGCT 198
QY 360 IleCysSerAspGlyAlaAspLeu-----GluLeuArgLeuAlaAspGlySerAsnAsn 377
Db 199 GTCACTACGCTCTGGAGAAATGAAGAGAACTGAGACTGCGGGTGTGTAAGACACAC 258
QY 378 CysSerGlyArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsn 397
Db 259 TGTACTGGGAGAGTGAACCTTAAGATCCATGACAAAGTGGGGCACAGTGTGCAGTAACGGC 318
QY 398 TrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPhe 417
Db 319 TGGAGCATGAATGAAGTGTCCCTGTTTCCAGCAGCTGGGATGCCCACTTCTATTAAA 378
QY 418 GlySerArgArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSer 437
Db 379 GCCTTGGATGGCTAACTCCAGCGCGCTCTGGATATATCTGGATGACAAAGTTTCT 438
QY 438 CysThrGlyAsnGluSerAlaLeuTrpAspCysThrTrpAspGlyLysAlaLysArgThr 457
Db 439 TGTACAGGGAATGAGTCACTCTTGGGACTCAAAAGATGATGGTGGGAAAGATAAC 498
QY 458 CysPheArgArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArg 477
Db 499 TGTACCATGAAAGATGCTGGAGTGACCTGCTCAGATGGATCAATTTGGAGATGAGA 558
QY 478 LeuVal---GlyAlaHisSerProCysTyArgLeuGluValValTyArgGlnGlyGlu 496
Db 559 CTGGTGAACAGTGGCGGCCACCGATCTTAGGAAGATAGATAAAGTTCCAGGAAAG 618
QY 497 TrpGlyThrValCysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGln 516
Db 619 TGGGGGACGGTGTGTGACGACAACTTACAGAAAGATGACCTCTCTCTGGGACTGCAA 678
QY 517 LeuGlyCysGlyLysProMethHisValPheGlyMetThrTyPheLysGluAlaSerGly 536
Db 679 CTGGATGGGAAGTGGCCATTAGTTTCTTGGCTCAGCTAAATTTGGAGCTGGTCTGGA 738
QY 537 ProIleTrpLeuAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGlu 556
Db 739 CCAATCTGGCTGATGACCTGGCATGCAATGGAATGAGTCACTCTCTGGGACTGCAA 798
QY 557 HisSerGlyTrpGlyLysHisAsnCysValHisArgLysAspValIleValThrCysSer 576
Db 799 CACCGGGGATGGGGAAGCATAACTGTGACCATGCTGAGGATGTGCGTGTGATTTGCTTA 858
QY 577 GlyAspAlaThrTrpGlyLeuArgLeuValGlyLysSerAsnArgCysSerGlyArgLeu 596
Db 859 GAGGAGCAGATCTGAGCCTGAGACTAGTGGATGGAGTGTCCAGATGTTCCAGAAAGATTG 918
QY 597 GluValTyPheGlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAla 616
Db 919 GAAGTGAGATTCAGAGGAATGGGGACCGGTGTGTGATGATGATAACTGGGATCTCCGGAT 978

```

Qy 617 AlaAlaValValCysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGly 636
Dy 979 GCTTCTGTGGTGTCAAGCAACTGGGATGCTCAACTGCCATCAGTGCATGGTGGAGTT 1038
Qy 637 AsnAlaSerThrGlyThrGlyLysIleThrLeuAspValSerCysAspGlyAspGlu 656
Dy 1039 AATGCCAGTGGAGTATCGGACAGATTTGGCTTGACAACTTTCATGCGAAGACATGAG 1098
Qy 657 SerAspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGlu 676
Dy 1099 GCAACTCTTTGGGATGTAAACACCAAGAGTGGGGAAGCAATTACTGTTCATCATAGAGAA 1158
Qy 677 AspValGlyValIleCysSerAspAlaSerAspMetGlnLeuArgLeuValGlyCysLeu 696
Dy 1159 CAGCTGTGTGACATGCTTCATGGACACATCTGGAACCTAGACTGTAGTGGAGGC 1218
Qy 697 SerArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAla 716
Dy 1219 AGTCGCTGTGCTGGCATTTGTGGAGGTGGAGATTCAGAAAGCTCAGCTGGGAAGATGTGAGC 1278
Qy 717 AsnGlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAla 736
Dy 1279 CGAGCTGGACACTGGCAGATGCGGATGTTGCGACACAGCTTGGATGTGGATCTGCG 1338
Qy 737 IleArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsn 756
Dy 1339 CTTCAAAACCCAGGCTAAGACTACTCTAAACCTGGGGCAACAATACGTGCTCTTCCT 1398
Qy 757 SerGlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGln 776
Dy 1399 GGATCTTGTAAATGGAATGAACACTACTTTTGGCAATCAAAACACTGCGAGTGGCGGC 1458
Qy 777 ThrAlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGluProArg 796
Dy 1459 CTTTCTCTGTGATATTTCGAAGAGCCAAAGTTACTCTCAGGCCACAGGGAAACCCAGA 1518
Qy 797 LeuValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrp 816
Dy 1519 CTGGTGGAGGAGAAATCCCATGCTCTGCTGTGGAGGTGAACACGGAGACGTGTGG 1578
Qy 817 ArgSerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeu 836
Dy 1579 GGCTCCGCTGTGATTTGACTTGTCTCTGGAAGCTGCACGTGTGTGTCGAGGGAATTA 1638
Qy 837 AsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeu 856
Dy 1639 CAATGTGGAACAGTCTCTATCTAGGGGAGGCACATTTTGGAGAAGAGTGGACAG 1698
Qy 857 ThrTrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIle 876
Dy 1699 ATCTGGGTGGAAGATTCAGTGTAGTGGGATGAGTCCCATCTTTCATATGCTCAGTG 1758
Qy 877 ValGlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArg 896
Dy 1759 GCGCCCCGCTGACAGAACTGTACCCACACAGAGGATGTGAGGATGTGCTGCTCAGCA 1818
Qy 897 TyrThrAspValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluLeuAsn 916
Dy 1819 TACATAGATATCTGCTGGCAGCGGCGGAGTCCCTCTGTGAGGGAAGAGTGGAGCTCAAG 1878
Qy 917 ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal 936
Dy 1879 ACACCTGGAGCCTGGGGTCCCTCTGTCAGTCTCATCTGGACATGGAAGATGCTCATGTC 1938
Qy 937 LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrlleGly 956
Dy 1939 TTATCTCAGCAGCTGAAGTGTGGGTGTCCTCAATCTATTCAGAAAGGAGACACATTTTGG 1998
Qy 957 GluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu 976
Dy 1999 AAAGGAGTGTGAGTCTGGATCAGATGTTCCACTGCTGGAACCTGAGGAACATATA 2058

Qy 977 AspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSer 996
Dy 2059 GGAGATTGCCCTCATGCTGCTGGTGGCGCAGCTGTCCTCCGAAGGACAGGTGGCCCTCT 2118
Qy 997 ValIleCysThrGlySerLeuThrGlnProLeuPheProCys-----LeuAlaAsnVal 1014
Dy 2119 GTCATCTGCTCAGAAACCAATCCAGACACATATTGGCATATTGCTTCTCCAGTCT 2178
Qy 1015 SerAspProTyrLeuSerAlaValProGluGlySerAlaLeuLeuLeuGluAspLys 1034
Dy 2179 CAACAACA---AGCTCTACAAATCCAAAGGAGAGTGAAGTTCCTGCTAGCAAGTGGC 2235
Qy 1035 ArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisasp 1054
Dy 2236 CAGCTTCGCTTGGTAGGTGGAGTGGCTCGCTGCGCTGGAAGAGTGGAGGTCTGCCACGAG 2295
Qy 1055 GlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValCys 1074
Dy 2296 GGCTCTTGGGGCACCGCTCTGTGATGACAAATGGGATATGACTGATGCCAATGTGGTGTGC 2355
Qy 1075 GlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGly 1094
Dy 2356 AAGCAGCTGGACTGTGGCTGGCAATTAACGCCACTGCTCTGCTTACTTGGGGAAGGA 2415
Qy 1095 SerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGln 1114
Dy 2416 GCAGGAGCTATCTGGCTAGACGAAGTCTATCTGCACCTGGGAAAGAGTCTCATATTGGCAG 2475
Qy 1115 CysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIle 1134
Dy 2476 TGCCATTCCATGAGTGGGGACGCCATACTGCAGGCCAACAAAGAGATGCGAGGTGTATC 2535
Qy 1135 CysSerGluPheThrAlaLeuArgLeuTyrSerGlnThrGluThrGluSerCysAlaGly 1154
Dy 2536 TGCTCCGAGTTCATGCTCTGAGGCTGACCAACGAGGCCACAAAGAAATGTCACAGT 2595
Qy 1155 ArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThr 1174
Dy 2596 CGCCTTGAAGTGTTTTACAATGGTACATGGGCGCAGTATTGGCAGTACCAATATGCTCCA 2655
Qy 1175 AlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeu 1194
Dy 2656 ACCACTGTGGGGGTGGTGGCTGCTGAGTGGCTGTGCAGACACAGGAGCTGTGAACCC 2715
Qy 1195 AlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLys 1214
Dy 2716 ATACCTTCAGACAGACACACCATCCAGGCCATGTGGGTAGATGCTGTCGAGTGTCCAAA 2775
Qy 1215 ThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerPro 1234
Dy 2776 GGAGTTGACACTTTGTGGCAGTGGCCCTCGTCACCTTGGAAACACAGACAGGCCAGCCC 2835
Qy 1235 ---AlaGluGluThrTrpIleThrCysGluAspArgIleArgValArgGlyGlyAspThr 1253
Dy 2836 TCTTCCAGAGTCTCTGGATCATCTGTGACAAACAAATAAGACTCCAGGAAGGCGATACA 2895
Qy 1254 GluCysSerGlyArgValGluIleThrHisAlaGlySerTrpGlyThrValCysAspAsp 1273
Dy 2896 GACTGTCTCGACGTGTGGAGATCTGGCACAAGGTTCTTGGGGAACAGTGTGTGATGAC 2955
Qy 1274 SerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeu 1293
Dy 2956 TCTTGGGATCTTAATGATGCTAAGTGTATGTAAGCAGTGTGGCTGTGGCAAGCTGTG 3015
Qy 1294 AlaAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMet 1313
Dy 3016 AAGGCACATAAAGAGACAGCATTTGGTCCAGGAACCTGGGCCCATATGCTCAATGAAAT 3075
Qy 1314 ArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSer 1333
Dy 3076 AAGTGTAGAGGAATGAGTCTTCCCTGTGGGATGCTCTCCCAACCCGTGGAGTGCACAGC 3135
Qy 1334 AspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeu 1353

polyA_site
BASE COUNT 933 a 1113 c 1442 g 987 t
ORIGIN

```

HVWRCPSSRGWGHCHHKDAGVICSEFLALRWSEDQOCACGWLVEFYINCTWGSVCVRH
PMEIDTSTICRQLCGDSDTSNVSVALREGPERPWVDRIQCRTDTSLWCQSPDPHN
YNCSGPKEEAYIWCADSRQIRLVGGGRCSRVEILLDOGSWGTICDDMDLDDARVVC
KOLQGDLATDVTSFGSGPGIWLDEYNTCKESHVWRPCSRGWGRHDCRHEDAGVTC
VICSGFRLAGDGDFCPSGRVEHSGEAMTPVSDGNFTLPQAIVCAELGCCKAVSVLG
HMPPRESQGVMAEEFRCDGBPELWSCPRVCPGCTCLHSGAAOVSCVSYTEVOLMK
NCTSCEGOVEMKISGRWRALCASHWLSANANVVCRQLCGSVAISTPRGPHLEVGDQ
ISTAQFHCBSFHLWSCPDPVTLAPGPCDSHGNTASVICGGNVTOVLPOCNDFLSQPAG
SAASESSPYCDSRQLRLLVGDCGGRVEILLDOGSWTICDDMDLDDARVVCROL
GCCEALNATGSAHFAGAGSGPIWLDLLNCTGKESHVWRPCSRGWGRHDCRHEDAGVTC
SEFLALRMYSDEQQCACGLLEVPWTGWSVCRSPMEDITVSVICROLRWRAERALSYYEDA
VGLRGSRPWNDDLQCRKWDYSLMOCPSPWKYSCSPKEEAYISCEGRRPKSCLNSTPA
AACTDREKLRLRGDSCESGRVEVWHNGSWGTICDDDSWSLAEEVYCQQLGCCQALEA
VRSAPFGPONGSIWLDEVQCGRESSLWDCVAEPWGSOCKHEEDARGVSCRVTLLP
TTTAGTRTNSLPGIFSLPVLCILIGSLLLFLVILVLTQLLRWAERALSYYEDALA
LAEAVYEELDYLLTKOEGLSGPDQMTDPDENYDDAEVYPGTPSPSQONEEVPEPE
KEDGVRSSQTGSGFLNFREAANPGEGESFWLLQKKKGADGYDDVELSALTGTSPVTF
"

```

4453

Alignment Scores:

Pred. No.:	9_9e-192	Length:	4475
Score:	3082.00	Matches:	613
Percent Similarity:	55.85%	Conservative:	208
Best Local Similarity:	41.70%	Mismatches:	495
Query Match:	37.87%	Indels:	154
DB:	4	Gaps:	23

US-09-759-130B-381 (1-1453) x BBWC11MR (1-4475)

QY	28	CysileLeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGlyThrAspLeuGlu	47
Db	58	TGNTGCTCTCTCTCTCGCC-----ACCATTGGTGCGTGTGAAGCTCTCGAG	102
QY	48	LeuArgLeuValAsnGlnGlyAspGlyProCysSerGlyThrValGluValLysPheGlnGly	67
Db	103	CTCAGGTGTAAGATGGAGTCCATCGCTGTCGAGGGGAGAGTGGAAAGTGAACACCAAGGA	162
QY	68	GlnTrpGlyThrValCysAspAspGlyTrpAsnThrAlaSerThrValValCysLys	87
Db	163	GAATGGGCGACAGTGGATGTTACAGTGGACATTTGAAGGATGCATCTGTAGTGTGCAGA	222
QY	88	GlnLeuGlyCysPropHeSerPheAlaMet-----PheArgPheGlyGlnAlaVal	104
Db	223	CAGCTGGGGTGTGGAGTGCATTTGTTCTCGAGGGGCTTATTTTGGCCAGGACTT	282
QY	105	ThrArgHisGlyLysIleTrpLeuAspValSerCysTrpGlyAsnGluSerAlaLeu	124
Db	283	-----GGCCCCATTTGGCTTTCTATACTTTCATGTGAAGGCAGACAGCTCAACTGTC	333
QY	125	TrpGluCysGlnHis-----ArgGluTrpGlySerHisAsnCysTyrHisGlyGlu	141
Db	334	AGTGACTGTGAGCATCTATAATATAAGACATTCGTAAATGATGGCTATAATCATGGTGGC	393
QY	142	AspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsn	161
Db	394	GATCCTGGAGTACTGCTCTCAGGA-----TTTGTGCGCTCTCGCTGGAGGGCAT	441
QY	162	AsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAsp	181
Db	442	GGACCTCTGCTCAGGGCGAGTAGAAGTGCATCTCTGGAGAAAGCTTGGATCCCAGTCTCAT	501
QY	182	AspGlyTrpAsnLeuAsnThrAlaAlaValValCysargGlnLeuGlyCysproSerSer	201
Db	502	GGGAACCTTCACACTTGCACCTGCCACATCATCTGTCAGAGTGGGTGTGGCGAAGCT	561
QY	202	PheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAspAsp	221
Db	562	GTGTCGTCTCTGGGACATGAGCTCTTCAGAGAGTCCAGTGCCCAGGCTGGGCTGCAAGAG	621
QY	222	IleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsn	241

Db 622 TTCAGTGTGAGGGGAGGAGCCTGAGCTCTGGGTCTGCCCCAGAGTGCCTGTCCAGGG 681
QY 242 HisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuGlu 261
Db 682 GGCAGGTGCACACAGCTGATCTGTCAGCTGTGTTGTTGTCAGCATACTCAGAAGTCGG 741
QY 262 LeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuIleGlnGly 281
Db 742 CTC---ATGACAACGGCTCTCTCAGTGTGAAGGCGAGTGGAGATGAACATTCTGGA 798
QY 282 ArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAspValValCysLys 301
Db 799 CAATGGAGAGCGCTCTGTGCTCCCTCAGTGTGAAGGCGAGTGGAGATGAACATTCTGCT 858
QY 302 GlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGlySer 321
Db 859 CAGCTGGGTGTGGAGTGGCATCTCCACCCCGGAGGACCATCTGGTGGGAAGAGT 918
QY 322 AspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCys 341
Db 919 GATCAGATCTAACAGCCGCTTCTCAGTGTCTGGGCTGAGTCTCTCTGTGGATTGT 978
QY 342 ArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCys 361
Db 979 CCTGTGACTGCCCTGGGTGCTGCTGACTGTCCATGGCAACACAGCCTCTGTGATCTGC 1038
QY 362 SerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArg 381
Db 1039 TCAGGAAC----- 1047
QY 382 ValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGlu 401
Db 1048 ---CAGATCCAGGTGCTTCCCGAG----- 1068
QY 402 GlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArg 421
Db 1069 -----TGCAACGACTCCGTG----- 1083
QY 422 AlalysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsn 441
Db 1084 TCTCACTACAGGCTCGCGGCTCA----- 1110
QY 442 GluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArgArg 461
Db 1110 ----- 1110
QY 462 SerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAla 481
Db 1111 GAGGACAGCGCCCTTACTGCTCAGACAGCAGG-----CAGTCCCGCTGGTGACGGG 1164
QY 482 HisSerProCysTyrGlyArgLeuGluValIlystYrGlnGlyGluTrpGlyThrValCys 501
Db 1165 GCGCGTCCCTGCGCGGAGAGTGGAGATCCTTGACAGGGCTCTCGGGGACCATCTGT 1224
QY 502 HisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLys 521
Db 1225 GATGACGGCTGGGACCTGGAGATGCCCGGTGTGTGACAGGAGCTGGGCTGTGGAGA 1284
QY 522 ProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAsp 541
Db 1285 GCCCTCAATGCCAGGGGTCTGCTCACTTCGGGCGAGGATCAGGGCCCATCTGTTGGAC 1344
QY 542 AspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGly 561
Db 1345 AACTTGAATGACAGACAGAGGAGTCCACAGTGTGGAGTGGCTTCCCGGGGCTGGGG 1404
QY 562 LysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrp 581
Db 1405 CAGCACAACTGCACACACAGAGGAGCGGGGGTCTATCTGCTCA-----GAGTTCCTG 1458
QY 582 GlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGln 601
|||||:|||||

Db 1459 GCCTCAGGATGGTGTGAGTGAGGACCACTGCTGGGTGGTGGAGTTTCTACAAT 1518
QY 602 GlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValValCys 621
Db 1519 GGCACCTGGGGCAGTGTCTGCCGTAAACCCATGGAACACATCACTGTCTCCAGCATGCG 1578
QY 622 SerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr--- 640
Db 1579 AGACAGCTTGGCTGTGGGGACAGT-----GGAACCCCACTCTCTCTGTGCT 1626
QY 641 -----GlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSer 657
Db 1627 CTTAGAAGAGGTTTATAGCCACACATGGGTGGATAGATCCAGTTCGGAACACTGACACC 1686
QY 658 AspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAsp 677
Db 1687 TCTCTCTGGCAGTGTCTCTGACCCCTTGAATTAACATCATCTCTCCAAAGGAGAA 1746
QY 678 ValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSer 697
Db 1747 GCCTATATCTGGTGTGCAGACAGC-----AGACAGATCCCGCTGGTGGATGGAGTGGT 1800
QY 698 ArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsn 717
Db 1801 CGCTGCTCTGGGAGTGGAGATCCCTTGACACGGGCTCTGGGGCACCATTGTGTATGAC 1860
QY 718 GlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIle 737
Db 1861 CGCTGGACCTGACGATGCCGTGTGTGCAAGCAGCTGGGCTGTGGAGAACCCCTG 1920
QY 738 ArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSer 757
Db 1921 GAGCCACTGCTCTCTCTTCTCGGCAGCGGATCAGGGCCCATCTGGCTGGATGAAGTG 1980
QY 758 GlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThr 777
Db 1981 AACTGTCAGAGAGAGAGTCCCAAGTATGGAGTGCCTTCTCTGGGATGGCGGCACAC 2040
QY 778 AlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeu 797
Db 2041 AACTGCAATCATCAAGAGATGACAGGATCATCTGCTCAGGATTT-----GTCCGCTG 2094
QY 798 ValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArg 817
Db 2095 GCTGGAGGAGATGGACCTGCTCAGGCGAGTAGAAGTGCAATCTGTGGAAGACCTGGACC 2154
QY 818 SerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsn 837
Db 2155 CAGTGTCTGATGAAACTTCACACTCCCACTCCCGAGGTCACTGTGCGAGAGCTGGGA 2214
QY 838 CysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThr 857
Db 2215 TGTGGCAAGCTGTCTGTCTCTGGGACACATGCCATTCAGAGAGTCCGATGGCCAGGTC 2274
QY 858 TrpAlaGlyLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleVal 877
Db 2275 TGGGCTCAAGAGTTCAGGTGTGTGGGGGAGCGCTCAGCTCTGCTGCTGCCCCAGAGTG 2334
QY 878 GlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyr 897
Db 2335 CCCTGTCCAGGAGGACATCTCTCCACAGTGGAGCTCTCAGGTTGTCTGTTCAGTGTAC 2394
QY 898 ThrAspValArgLeuVal---AsnGlyLysSerGlnCysAspGlyGlnValGluIleAsn 916
Db 2395 ACAGAACTCCAGTTATGAAACCGGACCTCTCAATGTGAGGGGCGAGTGGAGATGAAG 2454
QY 917 ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal 936
Db 2455 ATCTCTGGACGATGGAGAGCGCTGTGTGCTCCCACTGGAGTCTGGCCAATGCCAATGTT 2514
QY 937 LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGly 956
Db 2515 GTCTGTGTCAGTGGCTGTGGAGTGGCATCTCTCCACCCCGGAGGAGGACACATTGGTG 2574

Qy	1301	PheGlyGlnGlyThrGlyThrIleTrpLeuAspMetArgCysLysGlyAsnGluSer	1320
Db	3649	TTTGCCCTGGAAATGGGAGCATCGCTGGACAGAGTGCATCGGGGGCCGGAGATCC	3708
Qy	1321	PheLeuTrpAspCysHisAlaLysProTrpGlnSerAspCysGlyHisLysGluAsp	1340
Db	3709	TCCCTGTGGGACTGTGTTCGGAGCCCTGGGGGAGAGGACTCAAGCAGGAGGAT	3768
Qy	1341	AlaGlyValArgCysSerGly	1347
Db	3769	GCTGTGTGAGGTGCTCTGTGTGAAGCAACATTGCCACGACACAGCAGGACCAGA	3828
Qy	1348	-----GlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeuAlaLeuLeu	1364
Db	3829	ACAACCTCAATTCCTCCCTCCGTCATCTCCCTGCCTGGGGTCTCTGCCTATCCTG	3888
Qy	1365	SerIlePheGlyLeuLeuLeuValLeuPheLeuPheLeuThrTrpCysArg	1384
Db	3889	GGGTGCTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3945
Qy	1385	ValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArgGlySerLeuGlu	1404
Db	3946	GCAGAGCC-----AGACCTTATCCAGCTATGAGATGCTCTTGCT	3987
Qy	1405	GluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspProHisGlyThrArg	1424
Db	3988	GAAGCTGTGTATGAGGACCTCGATTACCTCTGCACACAGAAGGAGTCTGGCAGCCCA	4047
Qy	1425	ThrSer-----AspAspThrProAsnHisGlyCysGluAspAlaSerAspThrSerLeuLeu	1443
Db	4048	GATCAGATGACTGTGCTCCCTGATGAAATATGATGATGCTGGAAGAAGTACCAGTGCCT	4107
Qy	1444	GlyValLeuProAlaSerGluAlaThrLys	1453
Db	4108	GGAATCTCTCTCTCTCAGGGGAATGAG	4137
RESULT 11			
LOCUS	HS24243224	7656 bp	linear
DEFINITION	Homo sapiens mRNA for DMBT1 protein 8kb transcript variant 1		PRI 27-NOV-2000
ACCESSION	AJ243224		
VERSION	AJ243224.1	GI:5624921	
KEYWORDS	DMBT1 gene; DMBT1/8kb.1 protein.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 7656)		
JOURNAL	Mollenhauer, J., Holmskov, U., Wiemann, S., Krebs, I., Herbertz, S.,		
MEDLINE	Madsen, J., Kioschis, P., Coy, J.F. and Poustka, A.		
PUBMED	The genomic structure of the DMBT1 gene: evidence for a region with		
REFERENCE	susceptibility to genomic instability		
AUTHORS	Oncogene 18 (46), 6233-6240 (1999)		
TITLE	2 (bases 1 to 7656)		
JOURNAL	Mollenhauer, J.		
MEDLINE	Submitted (18-JUN-1999) Mollenhauer J., Molecular Genome Analysis.		
PUBMED	Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, 69120		
REFERENCE	Heidelberg, GERMANY		
AUTHORS	Location/Qualifiers		
TITLE	1..7656		
JOURNAL	/organism="Homo sapiens"		
MEDLINE	/db_xref="taxon:9606"		
PUBMED	/chromosome="10"		
REFERENCE	/map="10q25.3-q26.1"		
AUTHORS	/tissue_type="lung"		
TITLE	/dev_stage="adult"		
JOURNAL	1..7318		
REFERENCE	/gene="DMBT1"		
AUTHORS	1..106		
TITLE	5'UTR		


```
Db 4085 TCCCAACTGTGGCCATCATGAAGACGCTGTGTCTCATCTGCTCAGCTCCAGTCCAG 4144
Qy 1006 ProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaValProGluGly 1025
Db 4145 CCGACA---CCAGCCAGACACTTGGCCCACTCGCATGATCAACAGCAGGATCTGAA 4201
Qy 1026 SerAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1045
Db 4202 TCCAGTTTG-----GCCCTGAGGCTGGTGAATGGAGGTGCACAGGTGT 4243
Qy 1046 AlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrp 1065
Db 4244 CAGGCGCAGTGGAGGTCTATACGAGGCTCCTGGGCACTGCTGTGTGATGACTACTGTG 4303
Qy 1066 AspLeuSerAspAlaHisValValCysGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1085
Db 4304 GACCAATGATGCAATGTGTGTTCAGGAGCTGGCTGTGCTGGCCAGCTGACGCC 4363
Qy 1086 ThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCys 1105
Db 4364 CCAGGAAGTCCCGGTTTGGCCAGGTTTCAGACCCATTGCCCTGGATGATGGCGTGC 4423
Qy 1106 ThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCys 1125
Db 4424 TCAGGACATGATGCTATCTGTGGAGCTGCCGCCCAATGCTGGCTCTCCCAACATGT 4483
Qy 1126 ArgHisLeuGluAspAlaGlyValIleCys----- 1135
Db 4484 GGCCATCATGAAGACGCTGGTGTCTATCTGCTCAGCTTCCCACTCCAGCCAGCCAGCC 4543
Qy 1136 -----SerGluPheThr---AlaLeu 1141
Db 4544 CCAGACACTTGGCCAACTCTCGTGCATCAACAGCAGGATCTGAATCCACATTGGCCCTG 4603
Qy 1142 ArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsn 1161
Db 4604 AGACTGGTGAAT-----GGAGGTGACAGGTCTCGAGCCAGTGGAGTCCCTATACCAA 4657
Qy 1162 GlyThrTrpGlySerValGlyArgAsnIleThrAlaIleAlaGlyIleValCys 1181
Db 4658 GCCTCTGGGCGCAGCTGTGTGTGACTACTGGGACCAATGATGATGCCAAGCTGTGTCTG 4717
Qy 1182 ArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerTys----- 1199
Db 4718 AGGACAGTGGGCTGT-----GGCTGGGCCATGTGAGCCCAAGGAATGCCAGTTT 4768
Qy 1200 ---ThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSer 1218
Db 4769 GCCCAGGCTCAGGACCCATTTCTCTGTGATGATGTGCTGCTCAGGACACAGTCTTAC 4828
Qy 1219 IleTrpGlnCysLeuSerAlaProTrp----- 1227
Db 4829 CTGTGGAGCTGCCCCCACTAGTGGCTCTCCCACTGTGGCCATCATGAAGATGCT 4888
Qy 1228 -----GluArgArgIleSerSerProAlaGluGluThrTrpIleThr 1241
Db 4889 GGTGTCTATCTGCTCAGTGTCTCAGTCCCACTCAGCCCGCCAGCATACTTGGCTGACC 4948
Qy 1242 -----CysGluAspArgIleArgValArg----- 1249
Db 4949 ACCAACTTACCGGCAATTGACAGTAGATCTGAATCCAGTTTGGCTCTCAGGCTGTGAAT 5008
Qy 1250 GlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThr 1269
Db 5009 GGAGGTGAC---AGGTGTCGAGCGCGAGTGGAGGTGCTGTGATCGAGGCTCTCGGGAACC 5065
Qy 1270 ValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCys 1289
Db 5066 GTGTGTGATGACGCTGGGACACCAATGATCCCAATGTGGTCTCGAGGACAGCTGGGCTGT 5125
Qy 1290 GlySerAlaLeuAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1309
```

```
Db 5126 3GCTGGGCGCATCTGGCGCCAGGAAATCCCGGTTTGGCCAGGCTCAGGACCCATTGTC 5185
Qy 1310 LeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTyrAspCysHisAlaLysPro 1329
Db 5186 CTGATGATGTGGCTGTCTCAGGAAATGATGCTTACCTGTGGAGCTGCCCCCAAAAGGC 5245
Qy 1330 TrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSer 1349
Db 5246 TGGCTCACCCCAACTGTGGCCATCATCAAGACGCTGTGTCTATCTGCTCAGCCACCAAA 5305
Qy 1350 LeuLysSerLeuAsnAlaSerSerGlyHis 1359
Db 5306 ATAAATTTCTACTACACAGATTGGTGGCAT 5335
RESULT 12
LOCUS HSA243212 7683 bp mRNA linear PRI 17-SEP-1999
DEFINITION Homo sapiens mRNA for DMBT1 protein 8kb transcript variant 2
(DMBT1/8kb.2).
ACCESSION AJ243212
VERSION AJ243212.1 GI:5912463
KEYWORDS DMBT1 gene; DMBT1/8kb.2 protein.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 7683)
AUTHORS Holmskov,U., Mollenhauer,J., Madsen,J., Vitved,L., Gronlund,J.,
Tornoe,I., Kliem,A., Reid,K.B., Poustka,A. and Skjoldt,K.
Cloning of gp-340, a putative opsonin receptor for lung surfactant
protein D
Proc. Natl. Acad. Sci. U.S.A. 96 (19), 10794-10799 (1999)
MEDLINE 99415938
PUBMED 10485905
REFERENCE 2 (bases 1 to 7683)
AUTHORS Mollenhauer,J.
Direct Submission
TITLE Submitted (18-JUN-1999) Mollenhauer J., Molecular Genome Analysis,
Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, 69120
Heidelberg, GERMANY
FEATURES
Location/Qualifiers
1..7683
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/map="10q25.3-q26.1"
/tissue_type="trachea"
/dev_stage="adult"
1..7683
/gene="DMBT1"
1..106
/gene="DMBT1"
107..7345
/gene="DMBT1"
/function="putative tumour suppressor, role in immune
system and differentiation"
/notes="Sequence is an alternative splice form of the DMBT1
gene that is expressed in human adult trachea. Isoforms of
DMBT1 are identical to the collectin binding protein
gp-340. Full-length cDNA clone contains 1 bp deletions in
codons 100 and 1751, that were corrected by comparison
with the genomic exons"
/codon_start=1
/product="DMBT1/8kb.2 protein"
/protein_id="CAB56155.1"
/db_xref="GI:5912464"
/translation="MGISTVILEMCLLWGQVLSLGGWIPRTTDYASLIPSEVPLDITTV
AEGSPFPELTLESTVAEGSPISLESTLETTVAEGSLIPSESTLESTVAEGSDSGAL
RLVNGDRCQGRVEILYRGSHGAVCDSDNDANDVRCQLGCCWAMSAAPNAPFGQS
SGPITALDDVRCSHESYLSCPHNGHSHNGHGDGAGVICSAQPPQSLRPESMPVR
ISPPVPTGESSLALRLVNGDRCGRVEILYRGSMGTVCDDYDNDTNDANVCRLQ
CGWAMSAAPNAPFGQSGPIVLDDVRCSHESYLSCPHNGHSHNGHGDGAGVIC
AQSRLPTSPDPTWPTTASTAGPESLALRLVNGDRCQGRVEILYRGSMGTVCDDSW
```


QY 454 AlaLysArgThrCysPheArgSerAspAlaGlyValIleCysSer----- 469
 DB 2165 CTCCTCCACAACTGTGGCCCATCATGAAGTGTGCTCATCTGCTCAGCTGCCAGTCC 2224
 QY 469 ----- 469
 DB 2225 CGTGTGAGCCCGCCAGCCAGACAGTGTGTGACCATCATCGTTACTCTCATCGACAGTAGGA 2284
 QY 470 AspLysAlaAspLeuAspLeuValGlyAlaHisSerProCysTyrGlyArgLeu 489
 DB 2285 TCTGAATCCAGTTTGACCCCTGAGGCTGGTGAATGAAGTGCACAGGTGTACGGGCCGAGTA 2344
 QY 490 GluValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSerTrpArgAsn 509
 DB 2345 GAGTCTCATACCGAGGCTCTCTGGGCACCGTGTGTGATGACAGCTGGGATACCAATGAT 2404
 QY 510 AlaAlaValValCysLysGlnLeuGlyCysGlyLysProMetHisValPheGlyMetThr 529
 DB 2405 GCCAATGTAGTCTGACGAGCTGGGCTGTGGCTGGGCCACCGTGGGCCCGCCAGGAATGCC 2464
 QY 530 TyrPheLysGluAlaSerGlyProIleTrpLeuAspAspValSerCysIleGlyAsnGlu 549
 DB 2465 CGTTTGGCCAGGCTCAGGACCAATTTCTTGGATGATGTGCCCTGCTCAGGACACGAG 2524
 QY 550 SerAsnIleTrpAspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgGlu 569
 DB 2525 TCCTACCTGTGGAGCTGCCGCCCAATGGTGGCTCTCCACAACTGTGGCCATCATGAA 2584
 QY 570 AspValIleValThrCysSer----- 581
 DB 2585 GATGCTGTGTCTCATCTGCTCAGTTTCCAGTCCCGCGCCGACACCCAGTCCAGATACCTGG 2644
 QY 582 -----GlyLeuArgLeuValGly 587
 DB 2645 CCGACCTCATGATCATCAACAGCAGGACCTGAATCCAGCTTGGCCCTGAGGCTGGTGAAT 2704
 QY 588 GlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgTrpGlyThrVal 607
 DB 2705 GGAGTGCACAGGTCTCAGGCGCAGTGGAGTCTCTATACCGAGGCTCTCTGGGCACCGTG 2764
 QY 608 CysAspAspGlyTrpAsnSerLysAlaAlaValValCysSerGlnLeuAspCysPro 627
 DB 2765 TGTGATGATAGCTGGGACAGCAGTGCAGCCCAATGTGCTGCCGGCAGCTGGGCTGTGGC 2824
 QY 628 SerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyLysIleTrpLeu 647
 DB 2825 TGGCCACGTGACGCCCGCCAGGAATGCCCGTTTGGCCAGGCTTCCAGACCCATTTGCTCTG 2884
 QY 648 AspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgAsnSerGlyTrp 667
 DB 2885 GATGACGTGGCTGCTCAGGCTATGATGCTCTACCTGTGGAGTGCCTCCCAACAAATGGCTGG 2944
 QY 668 GlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerAspAla----- 685
 DB 2945 CTCCTCCATAACTGTACGACAGTGAAGAGCGTGTGCTCATCTGCTCAGCTGCCCACTCC 3004
 QY 685 ----- 685
 DB 3005 TGTGTGAGCCCGCCAGTCCAGACACATTTGCCGACCATCATCTTGGCTGCATCGACAGTAGGA 3064
 QY 686 -----SerAspMetGluLeuArgLeuValGlyGlySerArgCysAlaGlyLysVal 703
 DB 3065 TCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGTGCACAGGTGTACGGGCCGAGTG 3124
 QY 704 GluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsnIle 723
 DB 3125 GAGGTCTATACCAAGGCTCTCTGGGCGCCGCTGTGCGATGACAGCTGGGACACCAATGAT 3184
 QY 724 AlaGluValValCysArgGlnLeuGlyCysGlySerAlaIleArgValSerArgGluPro 743
 DB 3185 GCCAATGTGCTGCAGGCAACTGGGCTGTGGCTGGGCCCATGTGTCAGGCCCGCCAGGAATGCC 3244

QY 744 HisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGlyGlu 763
 DB 3245 CGGTTTGGTCAAGGCTCAGGACCCATGTCTGTGATGATGGCTGTCTCAGGACACAG 3304
 QY 764 AlaSerTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsnMet 783
 DB 3305 TCTTACCTGTGGAGCTGCCCCCAATGGCTGGCTCTCCCAACACTGTGGCCATAGTGAA 3364
 QY 784 GluAlaSerLeuIleCysSerAlaHisArg---GlnPro----- 795
 DB 3365 GACCTGGTGTCTCATCTGCTCAGCTTCCAGTCCCGGCCCAACACCTAGTCCAGACACTGG 3424
 QY 796 -----ArgLeuValGly 799
 DB 3425 CCAACCTCATGATCATCAACAGCAGGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAAT 3484
 QY 800 AlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerVal 819
 DB 3485 GGAGTGCACAGGTGTACGGCCGAGTGGAGTCTTATACCGAGGCTCTCTGGGACCCGTG 3544
 QY 820 CysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGly 839
 DB 3545 TGTGATGACTACTGGGACACCAATGATGCCAATGTGGCTTCCAGGACGCTGGGCTGTGGC 3604
 QY 840 AspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAla 859
 DB 3605 TGGGCCATGTGAGCCCGCCAGGAATGCCCGTTTGGCCAGGCTTCCAGGACCCATTTGCTCTG 3664
 QY 860 GluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHis 879
 DB 3665 GATGATGTGGCTGCTCAGGACATGAGTCTATCTGTGGAGCTGCCCCCACAATGGCTGG 3724
 QY 880 ProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAsp 899
 DB 3725 CTCCTCCCACTGTGGCCATCATGAAGACGCTGGTGTCTGCTCAGCTTCCAGTCC 3784
 QY 899 ----- 899
 DB 3785 CAGCCGACACCCAGCCAGACACTTGGCCAACTTCATCATCATCAACAGCAGGATCTCAA 3844
 QY 900 -----ValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIle 915
 DB 3845 TCCAGTTTGGCCCTCAGGCTGGTGAATGGAGGTGCACAGTGTCTCAGGCGCGAGTGGAGTC 3904
 QY 916 AsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArg 935
 DB 3905 CTATACCGAGGCTCTGGGCGACCGTGTGTGATGACTACTGGACACCAATGATGCCAAT 3964
 QY 936 ValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyLysTyrIle 955
 DB 3965 GTGTTTGCAGGACGCTGGGCTGGTGGCCAGTGCAGTCCAGCCCGAGGAAATGCCCGTTT 4024
 QY 956 GlyGluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeu 975
 DB 4025 GGCCAGGTTTCCAGGACCCATTTGCTCTGGATGATGTGGCTGTCTCAGGACATGAGTCCAT 4084
 QY 976 LeuAspAsnCysGlnMetThrValLeuGlyAlaProCysIleHisGlyAsnThrVal 995
 DB 4085 CTGTGGAGTCCCGCCCAATGGCTGGCTCTCCCACTCTGGCCCATCATGAAGACGCT 4144
 QY 996 SerValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSer 1015
 DB 4145 GGTGTCTATCTGCTCAGCTTCCAGTCCAGTCCCGGACGACACCCAGGACACTTGGCCA 4201
 QY 1016 AspProTrpLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArg 1035
 DB 4202 ACCTCATGATCATCAACAGCAGGATCTGAATCCAGTTTG-----GCC 4243
 QY 1036 LeuArgLeuAspAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGly 1055
 DB 4244 CTGAGGCTGGTGAATGGAGGTGACAGGTGTCTCAGGCCCGAGTGGAGGTCTCTATACGAGGC 4303
 QY 1056 PheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGln 1075

Db 4304 TCCTGGGGCCACCTGGTGTGATGACTACTGGGACACCAATGATGCCAATGTTGTTTCGAGG 4363
QY LysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySer 1095
Db 4364 CAGCTGGCTGTGGCTGGCCAGCTGAGCCAGGAAATGCCGGTTTGGCCAGGTTTCA 4423
QY GlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCys 1115
Db 4424 GGACCAATGCTGCTGATGTCGCTGCTCAGGACATGAGTCCCTATCTGTGGAGCTGC 4483
QY ProSerArgGlyTrpGlyGlnHisAspCysArgHisGlyGluAspAlaGlyValIleCys 1135
Db 4484 CCCACAATGGCTGGCTCTCCCAACAATGTGGCCATCATGAAGACGCTGCTCATCTCC 4543
QY SerGluPhe 1138
Db 4544 TCAGCTTTCAGTCCAGCCGACACCCAGCCAGACACTTGGCCAACTCTCGTCATCA 4603
QY 1139 -----ThrAlaLeuArgLeuTyrSerGluThrGluThrGluSer 1151
Db 4604 ACAGCAGGATCTGAATCCACTTTGGCCCTGAGACTGGTGAAT-----GGAGGTGACAGG 4657
QY CysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsn 1171
Db 4658 TGTCGAGCGGAGTGAGGTCTCTATACCAAGGCTCTGGGGACCGTGTGTGATGACTAC 4717
QY 1172 IleThrAlaIleAlaGlyLeuValCysArgGlnLeuGlyCysGlyGluAsnGlyVal 1191
Db 4718 TGGGACACCAATGATGCCAAGTGGTCTGCAGCAGCTGGGCTGT-----GGCTGG 4768
QY 1192 ValSerLeuAlaProLeuSerIlys-----ThrGlySerGlyPheMetTrpValAsp 1208
Db 4769 GCATGTCAGCCGACCAAGAAATGCCAGTGTGGCCAGGCTCAGGACCACTATCTCTGGAT 4828
QY 1209 AspIleGlnCysProIlyThrHisIleSerIleTrpGlnCysLeuSerAlaProTrp--- 1227
Db 4829 GATGTCGCTGCTCAGGACACGAGCCTTACCTGTGGAGTGGCCCAATGGCTGGCTC 4888
QY 1228 -----GluArgArgIle 1231
Db 4889 TCCACAACATGTGGCCATCATGAAGATGCTGTGTCTATCTGCTCAGTCTCAGTCCAG 4948
QY 1232 SerSerProAlaGluGluThrTrpIleThr-----GlyGlyAspThrGluCysSerGlyArgVal 1259
Db 5009 GAATCCAGTTTGGCTGAGGCTGGTGAATGAGGTGAC---AGGTGTCGAGGCCGAGTG 5065
QY 1260 GluIleTrpHisAlaGlySerTrpGlyThrValCysAspSerTrpAspLeuAlaGlu 1279
Db 5066 GAGGTCTCTATCGAGGCTCTCGGGAACCGTGTGTGATGACAGCTGGGACACCAATGAT 5125
QY 1280 AlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAla 1299
Db 5126 GCATGTCGCTGTCAGGACGCTGGCTGTGGCTGGCCATCTGCTGGCCCGCCAGGAATGCC 5185
QY 1300 SerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGlu 1319
Db 5186 CGTTTGGCCAGGCTCAGGACCCATTTGCTGGGTGATGTGCTGCTCAGGGAATGAG 5245
QY 1320 SerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGlu 1339
Db 5246 TCTTACCTGTGGAGCTGCGCCCAACAAAGGCTGGCTCAACCACTGGGCCATCACGAA 5305
QY 1340 AspAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHis 1359
Db 5306 GAGCTGGTGTCTATGCTCAGCCACCCCAATAAATTTCTACTACGACAGATGGTGGCAT 5365
RESULT 13
HSA297935

LOCUS HSA297935 7667 bp mRNA linear PRI 12-JAN-2002
DEFINITION Homo sapiens mRNA for DMBT1 8 kb transcript variant 2 (DMBT1 gene).
ACCESSION AJ297935
VERSION AJ297935.1 GI:14715230
KEYWORDS alternative splicing; deleted in malignant brain tumors 1; DMBT1 gene; DMBT1/8kb.2 protein.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
Mollenhauer, J., Herbertz, S., Helmke, B., Kollender, G., Krebs, I., Madsen, J., Holmskov, J., Sorger, K., Schmitt, L., Wiemann, S., Otto, H. F., Grone, H. J. and Poustka, A.
Deleted in Malignant Brain Tumors 1 is a versatile mucin-like molecule likely to play a differential role in digestive tract cancer
Cancer Res. 61 (24), 8880-8886 (2001)
JOURNAL 21621923
MEDLINE 11751412
PUBMED
REFERENCE 2 (bases 1 to 7667)
Mollenhauer, J.
AUTHORS Direct Submission
TITLE Submitted (09-NOV-2000) Mollenhauer J., Molecular Genome Analysis, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, 69120 Heidelberg, GERMANY
JOURNAL
FEATURES
Location/Qualifiers
source
1. 7667
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/map="10q25.3-q26.1"
/tissue_type="small intestine"
/dev_stage="adult"
88..7329
/gene="DMBT1"
88..7329
/gene="DMBT1"
/function="putative role in mucosal protection, immune response, wound healing"
/note="DMBT1 8kb transcript variant 2 deleted in malignant brain tumors 1 alternative splice variant, contains all exons except for exon 1"
/codon_start=1
/evidence=experimental
/product="DMBT1/8kb.2 protein"
/protein_id="CAC44122.1"
/db_xref="GI:14715231"
/translation="MGISTVILEVCLLWQVLSTGGWIPRTDYASLIPSEVLPDPTV
AEGSPFSESTLESTVAEGSPISLESTLESTVAEGSLIPSESTLESTVAEGSDSLAL
RLVNGDRCGRVEILYRGSGTVCDDSDWTNDANVYRQLCGGWAMAPGNFVGQ
SGPTLADDDVRCSGHESYLWSCPNGHSHNGHGEDAGVICSAAPQSTLTPESWVPR
ISPPVPTGSESSIALRLVNGDRCGRVEILYRGSGTVCDDSDWTNDANVYRQLG
CGWAMAPGNFVGQSGPTLADDDVRCSGHESYLWSCPNGHSHNGHGEDAGVIC
APQSRPTSPDTWPTSHASTAGPESLALRLVNGDRCGRVEILYRGSGTVCDDSD
WTNDANVYRQLCGGWAMAPGNFVGQSGPTLADDDVRCSGHESYLWSCPNGH
NCHSEDAGVICSAASRPTSPDTLPTITLTPASTGSESSIALRLVNGDRCGRVE
VLVYRGSGTVCDDSDWTNDANVYRQLCGGWAMAPGNFVGQSGPTLADDDVRC
ESYLWSCPNGHSHNGHGEDAGVICSGPESLALRLVNGDRCGRVEILYRGSG
TYCDDSDWTNDANVYRQLCGGWAMAPGNFVGQSGPTLADDDVRCSGHESYLW
NGLWLNCHGHEDAGVICSAASRPTSPDTLPTITLTPASTGSESSIALRLVNG
RCGRVEILYRGSGTVCDDSDWTNDANVYRQLCGGWAMAPGNFVGQSGPTLAD
DVRCSGHESYLWSCPNGHSHNGHGEDAGVICSVQSRPTSPDTWPTSHASTAG
ESSIALRLVNGDRCGRVEILYRGSGTVCDDSDWTNDANVYRQLCGGWAMAPGN
ARTGQSGPTLADDDVRCSGHESYLWSCPNGHSHNGHGEDAGVICSAASRPTSP
DTLPTITLTPASTGSESSIALRLVNGDRCGRVEILYRGSGTVCDDSDWTNDAN
VYRQLCGGWAMAPGNFVGQSGPTLADDDVRCSGHESYLWSCPNGHSHNGHGE
GVICSAASRPTSPDTWPTSHASTAGSESSIALRLVNGDRCGRVEILYRGSGT
CDDYWTNDANVYRQLCGGWAMAPGNFVGQSGPTLADDDVRCSGHESYLWSCP
NGLWLNCHGHEDAGVICSAASRPTSPDTWPTSHASTAGSESSIALRLVNGDRC
RVEILYRGSGTVCDDSDWTNDANVYRQLCGGWAMAPGNFVGQSGPTLADDDV
SGHESYLWSCPNGHSHNGHGEDAGVICSAASRPTSPDTWPTSHASTAGSESS

ARLVNGDRRCGRVLYRGSGVGTVDYDNDNDANVVCRLGCGWATSAPONAREG
 QSGSPVLDDVDSRGHESYLWSCPHNGHSHNGHEDAGVICSASQSPFTSPDTPW
 TSRTAGSESTLALRVNGDRRCRVEVLYOGSWGTVCDYDNDNDANVVCRLGCG
 GWAMSPAGNAGFQSGSPVLDDVDSRGHESYLWSCPHNGHSHNGHEDAGVICS
 AQOSTPRDPTWLTNLPAITVDSSESLALRVNGDRRCRVEVLYRGSGVGTVDYD
 NDNDANVVCRLGCGWAMSPAGNAGFQSGSPVLDDVDSRGHESYLWSCPHNGH
 HNGHEDAGVICSATQISTNTDWHPTTTTAPRESSNCGGFLFYASGTFSSPYA
 YTPNNAKCYWEIEVNSGYRINLGFSLNKLAEHNCSDYVEIFDGLNSLLIKTCN
 LYHGGTGWGTSWTIOAEVVCRLGCGRAVSALGNAYFGSGSPITLDDVCSGT
 RYHGGTGWGTSWTIOAEVVCRLGCGRAVSALGNAYFGSGSPITLDDVCSGT
 ESTLWCRNRGHSNCHREDAGVICSNHLSTPAPFLNITRPNIDYSCGGFLSOPS
 GFSSPFTGENTPNNAKCYWMDIEVQNNRYTVIFRDVLEGCCNYDIIEFDGPIESS
 PLIARCDGARGESSTSSNFMFSIRFSDHSITRRGEAYEYSSPNSDNLGLPNHM
 QASVRSYLSQISGLFASDLVISTWNGYIECRPOITNLVIFIPYSGCGTFKQADMT
 IDYSNFLTAAVSGGIKTRTDLRIHVSCRMLONTWDTMYIANDTHVANNTIQVEEV
 QYGNEDVNTSFTSSSFLYPTVSRPYVDLNOGLVVOAEILHSDAVLTLEVDTCVAP
 YSNDFTSLYDLIRSCVRDDTGPYSSPSLRARERFRAFLHFLNFPSPYLRCKMWV
 CRAIDFSSKCYRCVLRKSRKDVGSYQEKVDVVLGPIQLQOTAPRREPR"
 88..162
 /gene="DMBT1"
 163..7326
 /gene="DMBT1"
 /product="DMBT1/8kb.2 protein"
 /length="2166 g 1759. t"

sig_peptide

mat_peptide

BASE COUNT 1605 a 2137 c 2166 g 1759. t
 ORIGIN

Alignment Scores:
 Pred. No.: 6.39e-185 Length: 7667
 Score: 2982.00 Matches: 618
 Percent Similarity: 51.56% Conservative: 176
 Best Local Similarity: 40.13% Mismatches: 506
 Query Match: 36.64% Indels: 240
 DB: 9 Gaps: 23

US-09-759-130B-381 (1-1453) x HSA297935 (1-7667)

QY 44 ThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThrValGluVal 63
 DB 775 TCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGCGACAGGTGTCGAGGCCGAGTGGAGTGC 834
 QY 64 LysPheGlnGlyGlnTrpGlyThrValCysAspGlyTrpAsnThrThrAlaSerThr 83
 DB 835 CTATACCGAGGCTCCTGGGCGACCGTGTGTGATGACTACTGGGACACCAATGATGCCAAT 894
 QY 84 ValValCysLysGlnLeuGlyCysProPheSerPheAlaMet-----PheArgPhe 100
 DB 895 GTGCTCTGAGGAGCTGGGCTGGCTGGCCATGTCAGCCCCAGGAATGCCAGTTT 954
 QY 101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspValSerCysTyrGlyAsn 120
 DB 955 GGCCAGGCG-----TCAGGACCACTTGTCTCGATGATGCTGCTCAGGACAC 1005
 QY 121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140
 DB 1006 GAGTCTTACCTGTGGAGCTGCCCAATGAGTGGCTGGCTCACCACAACTGTGCCATAGT 1065
 QY 141 GluAspValGlyValAsnCysTyrGly----- 149
 DB 1066 GAGACGCTGGTGTCTGCTCAGCTCCCGAGTCCCGCGCGACACCCAGCCAGATACT 1125
 QY 150 -----GluAlaAsnLeuGlyLeuArgLeuVal 158
 DB 1126 TGGCCGACCTCAGATCATCAACAGCAGGACCTGAATCCAGTTTGGCCCTCAGGCTGGT 1185
 QY 159 AspGlyAsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThr 178
 DB 1186 AATGAGGTGACAGGTGTGAGGCGGAGTGGAGGTCTCTATACCGAGGCTCCTGGGGCACC 1245
 QY 179 IleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValCysArgGlnLeuGlyCys 198
 DB 1246 GTGTGTATGATAGCTGGGACACCACTGACGCCAATGTGTCTGCCGCGACCTGGGCTGT 1305
 QY 199 ProSerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrp 218

DB 1306 GGCTGGCCACGTCAGCCCGAGGAATGCGCGTTTGGCCAGGTTTCAGGACCCATTGTC 1365
 QY 219 LeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGly 238
 DB 1366 CTGATGACGTCGCGCTGCTCAGGCTATGAGTCTTACCTGTGGAGCTGCCCCACAAATGGC 1425
 QY 239 TrpGlyAsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyr----- 255
 DB 1426 TGGCTCTCCCACTACTGTCAGCACAGTGAAGACCGCTGGTGTCTCTGCTCAGCTGCCAC 1485
 QY 255 ----- 255
 DB 1486 TCCAGGTCGACGCCACGTCCTCCAGACACCGTTCGCCACCATCACCTTACCTGTCAGCAGTA 1545
 QY 256 ---AspSerSerAspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArg 274
 DB 1546 GGATCTGAATCCAGTTTGGCCCTCAGGCTGGTGAATGGAGGTGACAGGTGTTCAGGGCGGA 1605
 QY 275 ValGluLeuLysIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAla 294
 DB 1606 GTGAGGTCCTATACCGAGGCTCCTGGGACCCGCTGTGTGATGACAGCTGGGACACCAAT 1665
 QY 295 AlaAlaAspValValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeu 314
 DB 1666 GATGCCAATGTGTCTCAGGACGCTGGGCTGTGGCTGGCCATGTTGGCCCCCAGGAAT 1725
 QY 315 ProHisLeuGlnSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsn 334
 DB 1726 GCCCGTTTGGTCAGGCTCAGGACCCATGCTCTGTGATGACGCTGGCTGTCTAGGGAAT 1785
 QY 335 GluSerPheLeuTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGln 354
 DB 1786 GAGTCTTGTGGAGCTGCCCCACCAATGGCTGGCTCTCCCATAACTGTGGCCATAGT 1845
 QY 355 AsnAspValSerValIleCysSer---AspGlyAlaAspLeuGluLeuArgLeuAlaAsp 373
 DB 1846 GAAGACGCTGGTGTCTCAGGACCTGAAATCCAGTTTGGCCCTCAGGCTGGTGAAT 1905
 QY 374 GlySerAsnAsnCysSerGlyArgValGluValArgIleHisGluGlnTrpTrpIle 393
 DB 1906 GGAGTCAAGGTGTGAGGCGCGAGTGGAGTCTCTATACCGAGGCTTGTGGGGCACCGTG 1965
 QY 394 CysAspGlnAsnTrpLysAsnGlnAlaLeuValValCysLysGlnLeuGlyCysPro 413
 DB 1966 TGTGATGACAGCTGGGACACCAATGATGCCAATGTGTCTCAGGACGCTGGCTGGTGGC 2025
 QY 414 PheSerValPheGlySerArgAlaLysProSerAsnGluAlaArgAspIleTrpIle 433
 DB 2026 TGGGCCACGTCAGCCCGAGGAATGCCCGTTTGGTCAAGGCTCAGGACCCATTGTCTGT 2085
 QY 434 AsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLys 453
 DB 2086 GACATGTGGCTGCTCAGGACATGAGTCTTACCTGTGGAGCTGCCCCACAAATGGCTGG 2145
 QY 454 AlaLysArgThrCysPheArgArgSerAspAlaGlyValIleCysSer----- 469
 DB 2146 CTCCTCCACAACTGTGCCCATCATGAAGATGCTGTGTCTCATCTGCTCAGTCCAGTCC 2205
 QY 469 ----- 469
 DB 2206 CGGTGACGCCCGGCGGACACAGCTTGTGACCATCAGTTTACCTCCATCGACAGTAGGA 2265
 QY 470 AspLysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTyrGlyArgLeu 489
 DB 2266 TCTGAATCCAGTTTGACCCCTGAGGCTGGTGAATGGAAGTGCACAGTGTTCAGGCGCGGAGTA 2325
 QY 490 GluValLysTrpGlnGlyTrpGlyThrValCysHisAspArgTrpSerThrArgAsn 509
 DB 2326 GAGTCTCTATACCGAGGCTCCTGGGCGACCGTGTGTGATGACAGCTGGGATACCAATGAT 2385
 QY 510 AlaAlaValValCysLysGlnLeuGlyCysGlyLysProMetHisValPheMetThr 529

D	b	2386	GCCAATGTGCTGTCAGCAGCGCTGGCGCTGTGCTGGCGCCACGTCGGCGCCAGGAATGCC	2444
Q	y	530	TyrPheLysGluAlaSerGlyProIleTrpLeuAspAspValSerCysIleGlyAsnGlu	549
D	b	2446	CGGTTTGGCCAGGGCTCAGGACCCATTGTTCTGGATGATGTGGCTGCTCAGGACACGAG	2505
Q	y	550	SerAsnIleTrpAspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgGlu	569
D	b	2506	TCTTACCTGTGGAGCTGCCCCACAATGCTGGCTCTCCCACTGTGGCCATCATGAA	2565
Q	y	570	AspValIleValThrCysSer	581
D	b	2566	GATGCTGGTGTCTATCTGCTCAGTTTCCCACTCCCGCGCGACACCCAGTCCACATACTGG	2625
Q	y	582		587
D	b	2626	CCGACCTCACATGCATCACACGACGAGTCTGAATCCAGTTTGGCCCTCAGGCTGGTGAAT	2685
Q	y	588	GlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgTrpGlyThrVal	607
D	b	2686	GGAGTGCACAGGTGTGAGGCCGAGTGGAGGTCTTATACCGAGGCTCTCTGGGGCACCGTG	2745
Q	y	608	CysAspAspGlyTrpAsnSerLysAlaAlaValValCysSerGlnLeuAspCysPro	627
D	b	2746	TGTGATGATAGCTGGGACGACAGTACGCGCAATGTGTCTGCGCGCAGCTGGGTGTGGC	2805
Q	y	628	SerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyLysIleTrpLeu	647
D	b	2806	TGGGCCAGTTCAGCCCGCAGAAATGCCGGTTGGCCAGGTTTCAGGACCCATTGTCCTG	2865
Q	y	648	AspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgAsnSerGlyTrp	667
D	b	2866	GATGACGTGCGCTGCTCAGGCTATGAGTCTACCTGTGGAGCTGCCCCACAATGGCTGG	2925
Q	y	668	GlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerAspAla	685
D	b	2926	CTCTCCCATTAAGTCTCAGCACAGTGAAGACGCTGGTGTCTATGCTCAGCTGCCCACTCC	2985
Q	y	685		685
D	b	2986	TGTTGACGCCCGTAGTCCAGACACATTGGCCGACCATCAGCTTGCCTGCAATCGACAGTAGGA	3045
Q	y	686	SerAspMetGluLeuArgLeuValGlyLysSerArgCysAlaGlyLysVal	703
D	b	3046	TCTGAATCCAGTTTGGCCCTGAGGCTGGTGGATGATGAGGTGACAGGTGTACGGCCGAGTG	3105
Q	y	704	GluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsnIle	723
D	b	3106	GAGGTCTCTATACCAAGGCTCCTGGGGCACCGTGTGCGATGACAGTGGGACACCAATGAT	3165
Q	y	724	AlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGluPro	743
D	b	3166	GCCAATGTGCTGTGCAGGCAACTGGGCTGTGCTGGGCGCATGTGCAGCCCCAGGAATGCC	3225
Q	y	744	HisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGlyGlu	763
D	b	3226	CGGTTTGGTCAGGCTCAGGACCCATTGCTCTGGATGATGTGGCTGCTCAGGACACGAG	3285
Q	y	764	AlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsnMet	783
D	b	3286	TCTTACCTGTGGAGCTGCCCCACAATGGCTGGCTCTCCCACTGTGGCCATAGTGAA	3345
Q	y	784	GluAlaSerLeuIleCysSerAlaHisArg	795
D	b	3346	GACGCTGGTGTCTATCTGCTCAGTTCCTCCAGTCCCGGCCAACACCTAGTTCACAGACATTGG	3405
Q	y	796		799
D	b	3406	CCAACTTCACATGCATCAACAGCAGGATCTGAATCCAGTTTGGCCCTCAGGCTGGTGAAT	3465
Q	y	800	AlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerVal	819
D	b	3466	GGAGGTGACAGAGTGTACGGCCGAGTGGAGTCTTATACCGAGGCTCTCTGGGACCGGTG	3525

QY	820	CysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGly	839
DB	3526	TGTGAGTACTTGGCAGCACCAATGATGCCAAATGTGTTTGCAGGACAGCTGGGCTGTGGC	3585
QY	840	AspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTirPala	859
DB	3586	TGGGCCATGTACGCCCCAGGAATGCCCGGTTTGGCCAGGGTTTCAGGACCCATGTCTCCTG	3645
QY	860	GluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHis	879
DB	3646	GATGATGTGGCTGCCTCAGGACATGAGTCTCTATCTGTGGAGCTGCCCCCAACATGGCTGG	3705
QY	880	ProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAsp	899
DB	3706	CTCTCCCAACTGTGGCCATCATGAAGACGCTGGTGTCTATCTGCTCAGCTGCCAGTCC	3765
QY	899	-----	899
DB	3766	CAGCCGACACCCAGCCAGACACTTGGCCAACTCATCATCATCAACAGCAGGATCTGAA	3825
QY	900	-----ValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIle	915
DB	3826	TTCAGTTTGGCCCTGAGGCTGTGTAATGGAGGTGACAGGTGTCTCAGGCCCGAGTGGAGTCT	3885
QY	916	AsnValLeuGlyHisThrGlySerLeuCysAspThrHisIleTrpAspProGluLeuAspAlaArg	935
DB	3886	CTATACCGAGGCTCTGGGACCGGTGTGTGATGACTACTTGGGACACCAATGATGCCAAT	3945
QY	936	ValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIle	955
DB	3946	GTGGTTTGCAGGCAGCTGGGCTGTAGTGGGCCACGTCGACCCCGAGAAATGCCCGTTT	4005
QY	956	GlyGluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeu	975
DB	4006	GCCAGGGTTCAGGACCCATGTCTCTGGATGATGTGGCTGCTCAGACATGAGTCTCAT	4065
QY	976	LeuAspAsnCysGlnMetThrValLeuGlyAlaProCysIleHisGlyAsnThrVal	995
DB	4066	CTGTGGAGCTGCCCCACCAATGGCTGTTCTCCACAACTGTGCCCATCATGAAGAGCT	4125
QY	996	SerValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSer	1015
DB	4126	GGTGTCTATCTGCTCAGCTCTCCAGTCCAGCCGACACACTTGGCCCA	4182
QY	1016	AspProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArg	1035
DB	4183	ACCTCATATGCATCACAGCAGGATCTGAATCCAGTTTG-----GCC	4224
QY	1036	LeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGly	1055
DB	4225	CTGAGGCTGGTGAATGGAGGTGACAGGTGTTCAGGGCCGAGTGGAGGTCTCTATACCGAGGC	4284
QY	1056	PheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGln	1075
DB	4285	TCTTGGGGCAGCGGTGTGTATGACTACTGGGACCAACCAATGATGCCATGGTGTTCAGG	4344
QY	1076	LysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySer	1095
DB	4345	CAGCTGGGCTGTGGCTGGGCCACGTCAGCCCCAGGAAATGCCCGGTTTGGCCAGGGTTCA	4404
QY	1096	GlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGlnSerHisLeuTrpGlnCys	1115
DB	4405	GGACCATTTGCTGTGATGTGGCTGCTCAGGACATGAGTCCCTATCTGTGGAGCTGC	4464
QY	1116	ProSerArgGlyTrpGlyGlnHisAspCysArgHisLysLysAspAlaGlyValIleCys	1135
DB	4465	CCCCAATGGCTGGCTCTCCACACACTGTGGCCATCATGAAGACGCTGTGTCTCATCTGC	4524
QY	1135	-----	1135
DB	4525	TCAGTTCTCCAGTCCCGAGCGGACACCCAGCCAGACACTTGGCCAACTCTCGTGTGCATCA	4584

Qy 1136 -----SerGluPheThr-----AlaLeuArgLeuTyrSerGluThrGluThrGluSer 1151
 Db 4585 ACAGCAGGATGAATCACTACCTGGCCCTGAGACTGCTGAAT-----GGAGGTGACAGG 4638
 Qy 1152 CysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgAsn 1171
 Db 4639 TGTGAGCGCGAGTGGAGGCTTATACCAAGGCTCTGGGCGACCGGTGTGATGACTAC 4698
 Qy 1172 IleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyVal 1191
 Db 4699 TGGGACACCAATGATGCCACCTGCTGTCGAGCGAGCTGGGCTGT-----GGCTGG 4749
 Qy 1192 ValSerLeuAlaProLeuSerLys-----ThrGlySerGlyPheMetTrpValasp 1208
 Db 4750 GCCATGTCAGCCCCCAGGAATGCCAGTGTGGCCAGGCTCAGACCCCATTTGCTCGAT 4809
 Qy 1209 AspIleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrp--- 1227
 Db 4810 GATGTGCGCTGCTCAGGACGAGCTTACCTGTGGAGCTGCCCCCAATGGCTGGCTC 4869
 Qy 1228 -----GluArgArgIle 1231
 Db 4870 TCCCAACTGTGCCCATCATGAAGTGGTGTCTCATCTGCTCAGCTGCTCAGTCCAG 4929
 Qy 1232 SerSerProAlaGluThrTrpIleThr-----Cys 1242
 Db 4930 TCAACGCCAGCCAGCATACTTGGCTGACCAACCACTTACCGGCATTGACAGTAGGATCT 4989
 Qy 1243 GluAspArgIleArgValArg-----GlyCysAspThrGluCysSerGlyArgVal 1259
 Db 4990 GAATCCAGTTTGGCTCTGAGGCTGGTGAATGGAGGTGAC---AGGTCTCAGAGCCGAGTG 5046
 Qy 1260 GluIleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGlu 1279
 Db 5047 GAGTCTCTGATCGAGGCTCTCGGGAACCGTGTGTGATGACACTGGGACACCAATGAT 5106
 Qy 1280 AlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAla 1299
 Db 5107 GCCAATGTGTCTGTCAGGCGAGCTGGGCTGTGGTGGCCATGTGCGGCCCAAGAAATGCC 5166
 Qy 1300 SerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGlu 1319
 Db 5167 CGGTTTGGCCAGGCTCAGGACCCATTTCTGATGATGATGCTGCTCAGGGAATGAG 5226
 Qy 1320 SerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGlu 1339
 Db 5227 TCCTACCTGTGGAGCTGCCGCCACAAAGGCTGGCTCACCACCACTGTGGCCATCAGAA 5286
 Qy 1340 AspAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHis 1359
 Db 5287 GAGCTGTGTCTATCTGCTCAGCCACCAATAAATCTACTAGCAGATTTGCTGGCAT 5346

RESULT 14
 AF159456
 LOCUS AF159456
 DEFINITION Homo sapiens gp-340 variant protein (DMBT1) mRNA, complete cds.
 ACCESSION AF159456
 VERSION AF159456.1 GI:5733597
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 7686)
 Holmskov, U., Mollenhauer, J., Madsen, J., Vitved, L., Gronlund, J.,
 Tornoe, I., Kliem, A., Reid, K.B., Poustka, A. and Skjodt, K.
 Cloning of gp-340, a putative opsonin receptor for lung surfactant
 protein D
 Proc. Natl. Acad. Sci. U.S.A. 96 (1999):10794-10799 (1999)
 MEDLINE 99415938
 PUBMED 10485905
 REFERENCE 2 (bases 1 to 7686)
 Holmskov, U., Mollenhauer, J., Madsen, J., Vitved, L., Gronlund, J.,

Tornoe, I., Kliem, A., Reid, K.B.M., Foustka, A. and Skjodt, K.
 Direct Submission
 Submitted (16-JUN-1999) Department of Immunology & Microbiology,
 Institute of Medical Biology, University of Southern Denmark,
 Odense University, Winsloe Parken 21.1, Odense 5000, Denmark
 Location/Qualifiers
 1. 7686
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="10"
 /map="10q25.3-26.1"
 1. 7686
 /gene="DMBT1"
 /note="deleted in malignant brain tumor 1"
 107. 7348
 /gene="DMBT1"
 /note="scavenger receptor cysteine-rich protein SRCR;
 putative receptor for gp-3"
 /codon_start=1
 /product="gp-340 variant protein"
 /protein_id="A049696.1"
 /db_xref="GI:5733598"
 /translation="MGISTVILEMCLLWGOVLSTGWHIPRTTASYALIPSEVPLDQTV
 AEGSPFSESTLESTAEAGSPISLESTLESTVAGSLIPSESTLESTVAGSGLAL
 RLNGDRCQGRVEILYRGSGTVCDDSDTNDANVYVQLGCGWAMSAFNGAFGQ
 SGPIALDDVRCSESSYLWSCPHNGWLSHNGHEDAGVICSAAQPOSTLRPESWVR
 ISPPVTEGSESSILALRLVNGDRCGRVEILYRGSGTVCDDSDTNDANVYVQLG
 CGWAMSAFNGAFGQSGPIVLDVRCSESSYLWSCPHNGWLSHNGHEDAGVICS
 APOSTPSPDTPWPTSHASTAGPESLALRLVNGDRCGRVEILYRGSGTVCDDSD
 TSDANVYVQLGCGWAMSAFNGAFGQSGPIVLDVRCSESSYLWSCPHNGWLSH
 NCQHSDEGAVICSAAHWSPTDTPITLPASTVGSSESLALRLVNGDRCGRVE
 LYLRGSGTVCDDSDTNDANVYVQLGCGWAMSAFNGAFGQSGPIVLDVRCSE
 ESYLWSCPHNGWLSHNGHEDAGVICSPESSILALRLVNGDRCGRVEILYRGSG
 TVCDDSDTNDANVYVQLGCGWAMSAFNGAFGQSGPIVLDVRCSESSYLWSCP
 NGWLSHNGHEDAGVICSAAQSRTPRDTLSTLPSTVGSSESLALRLVNGSD
 RCGRVEILYRGSGTVCDDSDTNDANVYVQLGCGWAMSAFNGAFGQSGPIVLD
 DVRCSESSYLWSCPHNGWLSHNGHEDAGVICSQSRTPSPDTPWPTSHASTAGS
 ESSILALRLVNGDRCGRVEILYRGSGTVCDDSDTNDANVYVQLGCGWAMSAF
 ARFGQSGPIVLDVRCSESSYLWSCPHNGWLSHNGHEDAGVICSAAHWSPTSP
 DTLPITLPASTVGSSESLALRLVNGDRCGRVEILYRGSGTVCDDSDTNDANV
 CROGCGWAMSAFNGAFGQSGPIVLDVRCSESSYLWSCPHNGWLSHNGHEDAG
 VICSASQSRTPSPDTPWPTSHASTAGSSESLALRLVNGDRCGRVEILYRGSGT
 VTV
 CDDSDTNDANVYVQLGCGWAMSAFNGAFGQSGPIVLDVRCSESSYLWSCPHN
 GWSHNGHEDAGVICSASQSRTPSPDTPWPTSHASTAGSSESLALRLVNGDRC
 RVBLYRGSGTVCDDSDTNDANVYVQLGCGWAMSAFNGAFGQSGPIVLDVRC
 SESSYLWSCPHNGWLSHNGHEDAGVICSASQSRTPSPDTPWPTSHASTAGSESS
 ALRLVNGDRCGRVEILYRGSGTVCDDSDTNDANVYVQLGCGWAMSAFNGAFG
 QSGPIVLDVRCSESSYLWSCPHNGWLSHNGHEDAGVICSASQSRTPSPDTPW
 TSHASTAGSESSILALRLVNGDRCGRVEILYRGSGTVCDDSDTNDANVYVQLG
 CGWAMSAFNGAFGQSGPIVLDVRCSESSYLWSCPHNGWLSHNGHEDAGVICS
 AQSQSRTPDTPWPTSHASTAGSSESLALRLVNGDRCGRVEILYRGSGTVCDD
 SDTNDANVYVQLGCGWAMSAFNGAFGQSGPIVLDVRCSESSYLWSCPHNGWLS
 HNGHEDAGVICSATQINSTTDMWHTTTTARSSNGGFLFYASGTFSSPYPA
 YPNNAKCWEIEVNSGIRNLGFSNLKLEAHNCSEDFYEIFDGLSNLSLLGKICN
 DTROITSSNMTIHFSDTONTGLAWNSFSPDNLRLVNLNSSTLLGKICN
 IYHGWTGTVCDDSDTNDANVYVQLGCGWAMSAFNGAFGQSGPIVLDVRCSE
 ESTLWOCNRNRFNSHNCNHRDEAGVICSNHLSTPAPLNTIRPNTDYSCGGLS
 GDFSSPYPCNYNNACVMDIEVNNYRVTVIFRDVQLGCGNNTDYIEVDFGSPRS
 PLIARVCDGARGSTSSNEMSTFISDHSITRGRFAEYSSPNDSTNLCLPNHM
 QASVRSYLSLGSFASDLVSTWNGYECRQITPNLVIPTIPIYSCGTFKQDNVT
 IDYSNLTAAVSGGIKRRDRLRIHSCRMQLNTWDTMTYIANDTHVANTIQVEE
 YQNFDFNLTFTSSSFLPVTSPRYVLDNOLYVQAEILHSDAVLLEFVDTCAVP
 YSNDFTSLYDLIRSGCVRDITGYSPSSPLARFARFARFHLNRPSPVYLRCKMVV
 CRAYDPSSRCYRCVLRKRDVGSYQEKVDVLGPIQLQTPPRREPR"

BASE COUNT 1614 a 2141 c 2168 g 1763 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 8.66e-185 Length: 7686
 Score: 2980.00 Matches: 618
 Percent Similarity: 51.62% Conservative: 177
 Best Local Similarity: 40.13% Mismatches: 505
 Query Match: 36.62% Indels: 240

DB: 9 Gaps: 23
us-09-759-130b-381 (1-1453) x AF159456 (1-7686)
Qy 44 ThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThrValGluVal 63
Qy 794 TCAGATTGGCCCTGAGCTGGTGAATGAGCGCGACAGGTGTGCGAGCGCGAGTGGAGGTC 853
Qy 64 LysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrAlaSerThr 83
Qy 854 CTATACCGAGGCTCTGGGGACCGCTGTGTGATGACTACTGGGACACCAATGATGCCAAT 913
Qy 84 ValValCysLysGlnLeuGlyCysProPheSerPheAlaMet-----PheArgPhe 100
Qy 914 GTGGCTGCGAGCAGCTGGCTGTGGCTGGCCATGTCCAGGACCCAGGAAATGCCAGTTT 973
Qy 101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTrpGlyAsn 120
Qy 974 GCGCCAGGCG-----TCAGGACCCATGTCCCTGGATGATGTGGCTGCTCAGGACAC 1024
Qy 121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTrpHisGly 140
Qy 1025 GAGTCTTACCTGTGAGCTGCGCCACCAATGGCTGGCTCACCACCACTGTGGCCATAGT 1084
Qy 141 GluAspValGlyValAsnCysTrpGly----- 149
Qy 1085 GAAGACGCTGGTGTCTATCTGCTCAGCTCCCGACGCTCCCGCGCGACACCCAGCCAGATACT 1144
Qy 150 -----GluAlaAsnLeuGlyLeuArgLeuVal 158
Qy 1145 TGGCCGACCTCATGTCATCAACACGAGGACCTGTAATCAGTTTGGCCCTGAGGCTGGTG 1204
Qy 159 AspGlyAsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThr 178
Qy 1205 ATFGAGGTGACAGGTGTGAGCGCGCGAGTGGAGTCTCTATACCGAGGCTCCTGGGCGACC 1264
Qy 179 IleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCys 198
Qy 1265 GTGTGTGATGATAGTGGAGCACACCACTGACGCCAATGTGTGTGCGCGAGCTGGGCTGT 1324
Qy 199 ProSerSerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrp 218
Qy 1325 GGCTGGCGCACGTCAGCCCGGAGAAATGCCCGTTTGGCCAGGTTTCAGGACCCATTTGTC 1384
Qy 219 LeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGly 238
Qy 1385 CTGGATGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1444
Qy 239 TrpGlyAsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTrp----- 255
Qy 1445 TGGCTCTCCCAATACTGTCAGCACAGTGAAGACGCTGGTGTCTGCTGCTGCTGCTGCTGCT 1504
Qy 255 ----- 255
Qy 1505 TCTGGTGCAGCGCCAGTCCAGACAGGTGGCGACCATACCTTACCTGTCATCGACAGTA 1564
Qy 256 ---AspSerSerAspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArg 274
Qy 1565 GGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGAGGTTGACAGGTGTGTCAGGCGCGA 1624
Qy 275 ValGluLeuLysIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAla 294
Qy 1625 GTGGAGGTCTCTACCGAGGCTCTGGGCGACCGTGTGTGATGACAGTGGGACACCAAT 1684
Qy 295 AlaAlaAspValValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeu 314
Qy 1685 GATGCCAATGTGCTGCGAGGACGCTGGGCTGTGGCTGGGCGCATGTGGCCCCAGGAAT 1744
Qy 315 ProHisLeuGlnSerGlyAspValValTrpLeuAspGlyValSerCysSerGlyAsn 334
Qy 1745 GCCCGGTTTGGTCAGGCTCAGGACCAATGTCCTGGATGACGTGCGCTGCTCAGGGAAT 1804
Qy 335 GluSerPheLeuTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGln 354

Db 1805 GAGTCTTACTTGTGGAGCTGCCCGCACAAATGGCTGGCTCTCCCAATACTGGCCATAGT 1864
Qy 355 AsnAspValSerValIleCysSer---AspGlyAlaAspLeuGluLeuArgLeuAlaAsp 373
Db 1865 GAAGACGCTGGTGTCTATCTGCTCAGGACCTGAATCCAGCTTTGGCCCTGAGGCTGTGTAAT 1924
Qy 374 GlySerAsnAsnCysSerGlyArgValGluValArgIleHisGlnGlnTrpTrpThrIle 393
Db 1925 GGAGGTGACAGGTGTGACGGCGGAGTGCAGTCTATACCGAGGCTCTTGGGCGACCGTG 1984
Qy 394 CysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGlyCysPro 413
Db 1985 TGTGATGACAGCTGGGACACCAATGATGCCAATGTGGTCTCAGGACACTGGGCTGTGGC 2044
Qy 414 PheSerValPheGlySerArgArgAlaLysProSerAsnGluAlaAspIleTrpIle 433
Db 2045 TGGGCCATGTGACCCCGAGGAATGCCCGTTTGGTCAGGCTCAGGACCCATTTCTCTG 2104
Qy 434 AsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysThrTrpAspGlyLys 453
Db 2105 GATGATGTGGCTGCTCAGGACATGAGTCTTACCTGTGGAGTGCCTCCCAACAATGGCTGG 2164
Qy 454 AlaLysArgThrCysPheArgSerAspAlaGlyValIleCysSer----- 469
Db 2165 CTCCTCCCAACTGTGGCCATCATGAAGATGCTGTGTCTGCTCAGTCCCGCAGTCC 2224
Qy 469 ----- 469
Db 2225 CGGTCGACCGCCGCGACAGCAGTGTGCGACCATCAGCTTACCTCCATCGACAGTAGGA 2284
Qy 470 AspLysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTrpGlyArgLeu 489
Db 2285 TCTGAATCCAGTTTGCACCTGAGGCTGGTGAATGAAGTGACAGGTGTGACGGCGGAGTA 2344
Qy 490 GluValLysTrpGlnGlyGluTrpGlyThrValCysHisAspArgTrpSerThrArgAsn 509
Db 2345 GAGGTCTTATACCGAGGCTCTGGGCGACCGTGTGTGATGACAGCTGGGATACCAATGAT 2404
Qy 510 AlaAlaValValCysLysGlnLeuGlyCysGlyLysProMetHisValPheGlyMetThr 529
Db 2405 GCCAATGTGTGCGAGCAGCTGGCTGGCTGGCGCATGTCAGCCCGACAGAAATGCC 2464
Qy 530 TyrPheLysGluAlaSerGlyProIleTrpLeuAspValSerCysIleGlyAsnGlu 549
Db 2465 CGGTTTGGCCAGGCTCAGGACCCATTTCTGTGATGATGTGCGCTGCTCAGGACACGAG 2524
Qy 550 SerAsnIleTrpAspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgGlu 569
Db 2525 TCTTACCTGTGGAGTGTGCGCCCAATGGCTGGCTCTCCACAACTGTGGCCATCATGNA 2584
Qy 570 AspValIleValThrCysSer-----GlyAspAlaThrTrp 581
Db 2585 GATGCTGGTGTCTATCTGCTCAGTTTCCCGAGTCCCGCGCGACACCCAGTCCAGATACTGG 2644
Qy 582 -----GlyLeuArgLeuValGly 587
Db 2645 CGGACCTCACATGCATCAACAGCAGGATCTGAATCCAGTTTGGCCCTGAGGCTGTGTAAT 2704
Qy 588 GlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgTrpGlyThrVal 607
Db 2705 GGAGGTGACAGGTGTGAGGCGGAGTGGAGGTCTTATACCGAGGCTCTTGGGCGACCGTG 2764
Qy 608 CysAspAspGlyTrpAsnSerLysAlaAlaValValCysSerGlnLeuAspCysPro 627
Db 2765 TGTGATGATAGTGGGACACACAGTGCAGCCCAATGTGGTGTGCGCGAGCTGGGCTGTGGC 2824
Qy 628 SerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTrpGlyLysIleTrpLeu 647
Db 2825 TGGGCGACGCTCAGCCCGCAGGAATGCCCGTTTGGCCAGGTTTCAGGACCCCATTTGCTTG 2884
Qy 648 AspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgAsnSerGlyTrp 667

Db 2885 GATGACGTGGCTGCTCAGGCTATGATCTCTACCTGTGGAGCTGCCCCCAATGGCTGG 2944
Qy 668 GlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerAspAla----- 685
Db 2945 CTCTCCCATTAACGTGACGACAGTGAAGCGTGGTGTCTATCTGCTCAGCTGCCCACTCC 3004
Qy 685 ----- 685
Db 3005 TGGTCGACGCCAGTCCAGACACATTGCCGACCATCACTTGGCTGCTGCATCGACAGTAGGA 3064
Qy 686 -----SerAspMetGluLeuArgLeuValGlySerSerArgCysAlaGlyLysVal 703
Db 3065 TCTGAATTCCTAGTTGGCCCTGAGCTGGTGAATGGAGGTGACAGGTGTCAGGCGCGAGTG 3124
Qy 704 GluValAsnValGlnGlnAlaValGlyLeuCysAlaAsnGlyTrpGlyMetAsnIle 723
Db 3125 GAGTCTCTATACCAAGCTCTCTGGGACCGTGTGGCATGACAGCTGGGACACCAATGAT 3184
Qy 724 AlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGluPro 743
Db 3185 GCCAATGCTGCTGCAGGCAACCGGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3244
Qy 744 HisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGlyGlu 763
Db 3245 CGGTTTGGTCAGGCTCAGGACCATTTGCTCTGGATGATGCTGCTCAGGACACAG 3304
Qy 764 AlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsnMet 783
Db 3305 TCTATCCGCTGGAGCTGCCCCCAATGGCTGGCTCTCCCACTGTGGCCATAGTGAA 3364
Qy 784 GluAlaSerLeuIleCysSerAlaHisArg---GlnPro----- 795
Db 3365 GAGCTGTGTCTATCTGCTCAGCTTCCAGTCCCGGCAACACCTAGTCCAGACACTTGG 3424
Qy 796 -----ArgLeuValGly 799
Db 3425 CCAACTCATCATCATCAACAGCAGGATCTCAATPCCAGTTTGGCCCTGAGGCTGGTGAAT 3484
Qy 800 AlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerVal 819
Db 3485 GGAGGTACAGGTGTCAGGCCGAGTGGAGGTCTATACCGAGCTCTCTGGGCGACCGTG 3544
Qy 820 CysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGly 839
Db 3545 TGTGATGACTACTGGGACACCAATGATGATGCTGTTGCGAGCAGCTGGCTGTGGC 3604
Qy 840 AspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAla 859
Db 3605 TGGGCCATGTCAGCCGCCAGGAAATGCCCGTTGGCCAGGTTTCAGGACCCATTGTCTCTG 3664
Qy 860 GluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHis 879
Db 3665 GATGATGGCTGCTCAGGACATGATGCTCTATCTGTGGAGTGCCTCCACCAATGGCTGG 3724
Qy 880 ProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAsp 899
Db 3725 CTCTCCCAACAATGTGGCCATCATGAAGACGCTGGTGTCTCTGCTCAGCTTCCAGTCC 3784
Qy 899 ----- 899
Db 3785 CAGCCGACACCCAGCCAGACACTTGGCCCAACCTTCATCATCAACAGCAGGATCTGAA 3844
Qy 900 -----ValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIle 915
Db 3845 TCCAGTTTGGCTTGGAGCTGGTGAATGGAGGTGACAGGTGTACGGGCCGAGTGGAGGTC 3904
Qy 916 AsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArg 935
Db 3905 CTATACCGAGGCTCTGGGCGACCGTGTGTGATGACTACTGGGACACCAATGATGCCAAT 3964
Qy 936 ValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIle 955
Db 3965 GTGGTTGACAGGAGCTGGGCTGTGGCTGGGCGCACGCTCAGCCCGGCAAGTGGCCGGTTT 4024

Qy 956 GlyGluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeu 975
Db 4025 GSCCAGGGTTTCAGGACCCATTTGCTCTGGATGATGTGCTCTCAGGACATGAGTCTCTAT 4084
Qy 976 LeuAspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrVal 995
Db 4085 CTGTGGAGCTGCCCCCAATGGCTCTCCCACTCTCCCACTCTGTCGCATCATGAAGACGCT 4144
Qy 996 SerValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSer 1015
Db 4145 GGTGCTCATCTGCTCAGCTTCCAGTCCAGCCGACA---CCAGCCAGACACTTGGCCA 4201
Qy 1016 AspProTrpLeuSerLeuValProGluGlySerLeuAlaLeuIleCysLeuGluAspLysArg 1035
Db 4202 ACCTCATCATCATCAACAGCAGGATCTCAATCCAGTTG-----GCC 4243
Qy 1036 LeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGly 1055
Db 4244 CTGAGGCTGGTGAATGGAGGTGACAGGTGTCAGGCGCAGTGGAGTCTATATCCGAGGC 4303
Qy 1056 PheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGln 1075
Db 4304 TCTTGGGCGACCGTGTGTGATGACTACTGGACACCAATGATGCCAATGTGTTGAGG 4363
Qy 1076 LysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySer 1095
Db 4364 CAGCTGGCTGTGCTGGCGCACCTCAGCCCCAGGAAATGCCGTTTGGCCAGGTTCA 4423
Qy 1096 GlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCys 1115
Db 4424 GGACCCATTTGCTTGGATGATGTGCTGCTCAGGACATGATGCTCTATCTGAGGCTGC 4483
Qy 1116 ProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCys 1135
Db 4484 CCCCAATGGCTGGCTCTCCCACTGTGCCATCATGAAGACGCTGGTGTCTATCTGC 4543
Qy 1135 ----- 1135
Db 4544 TCAGTTTCCAGTCCCGCCAGCACACCCAGCCAGACACTTGGCCCAACCTCTCGTGATCA 4603
Qy 1136 -----SerGluPheThr---AlaLeuArgLeuTrpSerGluThrGluThrGluSer 1151
Db 4604 ACAGCAGGATCTGAATCCACTTTGGCCCTGAGACTGTGTAAT-----GGAGGTGACAGG 4657
Qy 1152 CysAlaGlyArgLeuGluValPheTrpAsnGlyThrTrpGlySerValGlyArgArgAsn 1171
Db 4658 TGTGAGGCGGAGTGGAGGTCTTATACCAAGGCTCTCTGGGCGACCGTGTGTGATGACTAC 4717
Qy 1172 IleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyVal 1191
Db 4718 TGGCACACCAATGATGCCAAGCTGTGTCAGGACGCTGGGCTGT-----GGCTGG 4768
Qy 1192 ValSerLeuAlaProLeuSerLys-----ThrGlySerGlyPheMetTrpValAsp 1208
Db 4769 GCCATGTAGCAGCCCGAGGAAATGCCAGTTTGGCCAGGCTCAGGACCATTTGCTCGAT 4828
Qy 1209 AspIleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrp--- 1227
Db 4829 GATGTGCTGCTCAGGACACGAGTCTTACCTGTGGAGCTGCCGCCCAACAAATGGCTGGCTC 4889
Qy 1228 -----GluArgArgIle 1231
Db 4889 TCCCACTACTGGCCATCATGAAGATGCTGTGTCATCTCTCAGTGTCTCAGTCCAG 4948
Qy 1232 SerSerProAlaGluGluThrTrpIleThr-----Cys 1242
Db 4949 TCACGCCCGCAGCCAGATCTTGGCTGACCACTTACCGGCATTCACAGTAGTAGGATCT 5008
Qy 1243 GluAspArgIleArgValArg-----GlyGlyAspThrGluCysSerGlyArgVal 1259
Db 5009 GAATCCAGTTTGGCTCTGAGGCTGGTGAATGGAGGTGAC---AGGTGTGCGAGCGGAGTG 5065

[illegible]

Ds 78852 TAAATAATATATATAATAAACAATATATATAATAAACAATCTATATAAACCATTATAT 78793
Qy 549 ----- 549
Ds 78792 ATAAATATATAATATATATATAAACAATATACATATAAATATATATATATATATAT 78733
Qy 549 ----- 549
Ds 78732 AAAAACCTAAAAGTCAATATATGATCTACCTTTTATGCTGAATCACCCTACAGCAACTGT 78673
Qy 549 ----- 549
Ds 78672 TTTCCACCTAAGACTCTCTCTCAGAGTGAATATGTTCTACAAAGCAGTGAATATATGTT 78613
Qy 550 -----SerAsnIleTTPAspCysGluHisSerGlyTTPGlyLysHis-- 563
Ds 78612 TTTCCCTTATATTTAGAGTAGTCTCATTTTAATGAAGAAGTTCTATCTGTTCTAGACACCT 78553
Qy 563 ----- 563
Ds 78552 TTTTCAGGAATCTCAATTTTGTATAGAGATTCCTACAAGATATATATGTTTATTTCTCCCA 78493
Qy 563 ----- 563
Ds 78492 CCCTCCCAACACACACACAAAATATATAGCTCTCTAAACGTCAGATATCTAAATAAAC 78433
Qy 564 -----AsnCysValHisArgG1 569
Ds 78432 CCATTTAAAAAATTTCTTTTATGTTTATGTTAAAGTGTGAAGCTGATCTACCAACC 78373
Qy 569 uAspValIleValThrCysSerGlyAspAlaThrTTPGlyLeuArgLeuValGlyGlyse 589
Ds 78372 TCATGATATATGTT-----TTGAACCTCTGCCAATATAC 78337
Qy 589 rAsnArg-----Cys----- 592
Ds 78336 AAACAGATCTTGACTCAATGTTTCAAAACCCCTGTCAGCTGAGCTCCAAAGACACAGATA 78277
Qy 593 -----SerGlyArgLeuGluValTyrPheGlnGlyArgTTP----- 604
Ds 78276 CAATAGGCTTAAATGGAATATACATCAACTCTATTTCTGCTGAATGGAGTTTAA 78217
Qy 604 ----- 604
Ds 78216 TTTAGAGATCATCTATATCTACACAGATATATCAAAATAACCAATGATGATAAAT 78157
Qy 605 -----GlyThrValCysAspAspGlyTTP----- 612
Ds 78156 TCATTCAGCTGTTTATAAGTGGCAGGAAGTGGAGCTGA-AGTGATAGGAAGGACAGAC 78098
Qy 612 ----- 612
Ds 78097 TTTCCAGGCTCTAGACCAGGTTAATATATATATTTCTTCTTCAAGGCCAATAGGAAG 78038
Qy 612 ----- 612
Ds 78037 TAACCTTGCTCAGGTTTGATCTGGTCAAGTCAGAGTTGATTTTCTGACATGTCACATCTG 77978
Qy 612 ----- 612
Ds 77977 TCAACATTAGGAGTTAATAAACCTATTAAATATAATTTAGTATATAAAGTAGAAGAAATAG 77918
Qy 613 -----AsnSerLys 615
Ds 77917 TATGATATCAGATAGACATTTAATAATGTGAATTTTATTTTAACTAAACAATAGAAA 77858
Qy 615 sAlaAlaAlaValValCysSer-----GlnLeuAspCys----- 626
Ds 77857 AGAGCTGATACCATTTATAGTTTAAATAGATAAAGAAGTATTCATTTGTATATATAT 77798
Qy 626 ----- 626
Ds 77797 TAGATATATTTCTCAAAATTAGGCAATATTTATTTGTGTTTTCACAGACTATGCCACGCA 77738

Qy 626 ----- 626
Ds 77737 CACTCAAAAGCTACCAAACTGCCCTAAAAGATAAAGTAACCTTTCTAGAAATGAGATTCA 77678
Qy 626 ----- 626
Ds 77677 TTTCAATTTTATGTCATCTACTGCTAAAAAATTTCTCTCAAAAGCTGACTCTTTCAATAGGCT 77618
Qy 626 ----- 626
Ds 77617 CAGAATTAATTTCTGATTAATTTCTGCTTTCTCTCTTTTAAAAAAGTTTTTATTTTCCA 77558
Qy 626 ----- 626
Ds 77557 TAGGTTTTTAGGGAACAGATGGTAATAGTATACATGAGTAAGTTCCTTTAGTTTGTATTTGT 77498
Qy 627 -----ProSerSer----- 629
Ds 77497 GAGATTTTGTGATGCACTCATCACCAAGCAGTATACACTGAACCCAATTTGTAGTCTTTTC 77438
Qy 630 -----IleIleGlyMe 633
Ds 77437 TTTCTCAAAACCCCTCGTACCTTTTCCCTGAGTCTCTCAAAAGTCCATTTGATTTCTTAT 77378
Qy 633 tGlyLeuGlyAsnAlaSerThrGly----- 641
Ds 77377 CTCCTTGTGATCGTCATGCTTAGCCCTTGTAGCCCTTATGAGTGAAGACATACATATTTGGT 77318
Qy 641 ----- 641
Ds 77317 TTTCCATTTCTGATTTACTTTACTTTAGTAATAATAGTCTTCAATTTCCATCTGTTGCTGTG 77258
Qy 642 -----TyrGlyLysIleTrpLeuAspVal---SerCys----- 652
Ds 77257 AATGCCATTAATTCATCTCTTTATGCTGAGTAGTATTCATCATGCTGTATATATATA 77198
Qy 652 ----- 652
Ds 77197 CACACATACATGTGTATACGCATATACACATACATGTATGTATAGCGGTATACACA 77138
Qy 652 ----- 652
Ds 77137 CATACATGTATATACGTGTATACATATATATATATATATATATGTATATGTATACACACA 77078
Qy 653 -----AspGly----- 654
Ds 77077 CACACACACACACACACACACAGTTTCTTTATCCACTTTGTTGATTTGAGGCAATTTGG 77018
Qy 654 ----- 654
Ds 77017 CTGGTTCCATATTTTGCATTTGCGATTTATGCTCTATATAACATGCGTGTGCAAGTATC 76958
Qy 655 -----AspGluSerAspLeuTrpSer----- 661
Ds 76957 TTTTTCAGATAATAGCTATTTTCTCTCTGGGTAGATACCTAGTAGTGGGATTTGCTGGATC 76898
Qy 661 ----- 661
Ds 76897 AAATGGTAGCTCTGCTTTTATGTTTCTTTAAAGAAATCTCCACTGTTTTCATACGCGTTGTA 76838
Qy 662 -----CysArgAsn----- 664
Ds 76837 CTAGTTACATTTCCCAACAGCAGTGTAGAAGTTCCTTTCCACTGCATCCATGAAACT 76778
Qy 665 -----SerGly----- 666
Ds 76777 CTATTTCTTTTGTATTTTGTGTTTATGCTTCTTCGAGGAGTAAAGTGGCAATTTGCATT 76718
Qy 666 ----- 666
Ds 76717 GTGGTTTTGTGATTTGCTGATTTTCTGATCATTTAGTATGATGAGCATTTTTCATTTAAAAA 76658

Db 74498 GACATCAATTCCAAAAAGCCCTAAGAGGCCAATATAGATGAACACTGAACAATTTCCACT 74439
Qy 790 ----- 790
Db 74438 AAGCTTAGCTACAGAGAGTCATCAATAAATTTAAATAAGTAGTATGTCACATTAGAA 74379
Qy 790 ----- 790
Db 74378 AAAGGAGATCATTTCAAGTTGAGAAAGAAATGGGAGGTAAAGATGTTGGGATAATAGCA 74319
Qy 790 ----- 790
Db 74318 TAGACTATTCCAAAAAGCTGTATTTAAGGGGAAGAAATAAATAAAGTGATGACTAAAA 74259
Qy 790 ----- 790
Db 74258 AGGAATATAATCTTCAGGGGAGTTACTCAATTTTTTTTATTTAGTAGAGATCTTAGAGTA 74199
Qy 790 ----- 790
Db 74198 GTGCTGTTCAAAAGAACATCACCTTCTGTGATGAAGTAAATAATATATGCTATTAGCTGATG 74139
Qy 790 ----- 790
Db 74138 TTAGAGGCTATAGCAAAAAATAGAACTAGAAATAGAAATCAGAGTGTTTATGGACAGA 74079
Qy 790 ----- 790
Db 74078 ATGAATTTCCAAATAAATGACTGGGTAGGCCCTTACTGAAAGGTGAAACATAGACAA 74019
Qy 790 ----- 790
Db 74018 GACTCGAAGGTCATGGGAGAGTCAGTTATATGCATATACGAGAAGGAAATAACAATAG 73959
Qy 790 ----- 790
Db 73958 AGATACAGCCAGTTCAAAGCCTTGGGGCAAGATTATGCCGTGATGTTTGAAGAAGAGA 73899
Qy 790 ----- 790
Db 73898 AAGGAATCCANTGACGTGAAGACAGTAACTAAACATCAGACCCTGGCACCACAAAGGG 73839
Qy 790 ----- 790
Db 73838 GTAATGAGCAGGGGAAGGTCACCTCAGGGACAAGAGTCTTCGTAGGAGTTTAGATTTT 73779
Qy 790 ----- 790
Db 73778 TACTCTGAGTGAATATGGAGTCATGGGAGGATTTGGGGCAGAACTACTGATTTAACTGA 73719
Qy 790 ----- 790
Db 73718 TACATCATTTTAAATAACCTCTCTGGATTATTTTATTTTAAAGANTGAACATAGAGGA 73659
Qy 790 ----- 790
Db 73658 ATAATTATAGAAACAGGGAATCTCTAGGAATCTATTATAGTAAGTCACTCAGGCAAGTAATGAT 73599
Qy 790 ----- 790
Db 73598 GGTGGCTCAATCAAGGCCCTAGGGTTGAAGGGGTGAGAAGTTTGGGGATATGGGCATAC 73539
Qy 790 ----- 790
Db 73538 TTTGAAATACAGAAACAAACCTTCTGAAAGGGAGATATAGTATGTAATGGAAGTCA 73479
Qy 790 ----- 790
Db 73478 AGAATTATTTCAAAATATTGCAATCTCTGGAATGATGGAGTGTCCATCAATAGAAATAGG 73419
Qy 790 ----- 790
Db 73418 AAGTCTGAGCCTTAGATAGATGTTTGAAGAGGACTCCAAAGAATTCACATTTAGATATCT 73359

Qy 790 ----- 790
Db 73358 ATAAGAAATCCAAATGGAATGCTAAGTAGAAATTTGGATAAATGAGTCAGGAGTTTGGAA 73299
Qy 790 ----- 790
Db 73298 AAATGCCTGGTGTGGAGACAGATAACATACATTACTAATATAAGTAGCTTTGCTGCAAG 73239
Qy 790 ----- 790
Db 73238 AATAGAGTCAGAAATAATGGGAGTATTTTGTGTTTAAATTTACTCATTTAGATGG 73179
Qy 790 ----- 790
Db 73178 ATTGCTGGATCTAACCAAGTAGAAGTCTTTCTATCTTATATTACACCATCTTCTT 73119
Qy 790 ----- 790
Db 73118 AAGATGACTACAAATCTTCACGTGTGTTTAAATGCTGAGTTACTTAGTATTTTTTAAT 73059
Qy 790 ----- 790
Db 73058 GTAACAAATTTGTCCAATTCATTTCTTCGATCAGTTCTAAATCTAGGATTTGGTGGTC 72999
Qy 790 ----- 790
Db 72998 ACCAAGACATAGAAATAGGTATGTAAATCCAGATCTGTGAATTTAAGTCTGACTGTAA 72939
Qy 790 ----- 790
Db 72938 TCTTGTCTCTGAGCCCCCTACGAATGGGATGTTCTTAATGATGAATGGAAAAATGTC 72879
Qy 790 ----- 790
Db 72878 TGGGTTAGATATGGAGACTCGAGACAGGAGTAGATATTGGAACCCCTAAACCTAACC 72819
Qy 791 -----AlaHisArgGlnProArgLeuValGlyAlaAspMetProCy 804
|||||
Db 72818 TTTGAAATATTATTTTATAGCCACAGCCAGCGCTGTTGGAGCTGATATGCCCTG 72759
Qy 804 sSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSerAspH 824
|||||
Db 72758 CTCTGGAGCTGTTGAAGTGAACATGCAGACATGCGCGCTCTGTCTGTGATTTCTGATTT 72699
Qy 824 eSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIleSerLe 844
|||||
Db 72698 CTCCTTCATCTGCCAATGCTGTCGAGAGAATTAACCTGTGGAGATGCCATATCTCT 72639
Qy 844 uSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPheGlnCy 864
|||||
Db 72638 TTCCTGGGAGATCACCTTTGAAAAAGGAATGGTCACTTTGGGCGGAAAAAGTTCCAGTG 72579
Qy 864 sGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAspThrCy 884
|||||
Db 72578 TGAAGGAGTGAACCTCACCTTGCATTTATGCCCTTTCAACATCCCGAAGACACTG 72519
Qy 884 sIleHisSerArgGluValGlyValValLysSerArg 896
|||||
Db 72518 TATCCACAGCAGAGAAAGTTGGAGTGTCTGTCTCCCG-TGAGATTTTAAATAATCTTAAT 72460
Qy 896 ----- 896
Db 72459 TGGCTGGGCTAGAGATGGGAGTTTGGGCTGGAGCTGATAAGTCATGATGTCCTCTCT 72400
Qy 897 -----TyrThrAspValArgLeuValAsnGlyLysSerGlnCy 909
|||||
Db 72399 ATGCCTTGTCTCTCTCTAGGATATACAGATGTCGAGCTTGTGAATGGCAATCCCACTG 72340
Qy 909 sAspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTr 929
|||||
Db 72339 TGACGGGCAAGTGGAGATCAACGTCGTTGGACACTGGGGCTCACTGTGTGACACCCACTG 72280

QY 929 paspproGluuAspAlaAArgValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerTh 949
|||||
Db 72279 GGACCCAGAGATGCCCGTGTCTATGCAGACAGCTCAGCTGTGGAGCTGCTCTCAAC 72220
QY 949 rThrGlyGlyLysTyrIleGlyGluArgSerValArgValTrpGlyHisArgPheHisCy 969
|||||
Db 72219 CACAGGAGGAATATATTGGAGAAAGAGTGTGTGTGGGGACACAGGTTTCATGT 72160
QY 969 sLeuGlyAsnGlnSerLeuLeuAspAsnGlnMetThrValLeuGlyAlaProProCy 989
|||||
Db 72159 CITAGGGAATGAGTCACTCTCGATAACTGTCAATGACAGTCTTGGAGCACCCTCCG 72100
QY 989 sIleHisGlyAsnThrValSerValIleCysThr 1000
|||||
Db 72099 TATCCATGGAATACTGTCTGTGTATCTGCACAGGTAAGAGAATAGTCTTATGGTTAA 72040
QY 1000 1000
Db 72039 CTACTGTCAATATAGCCAGACATGTTAGAGGAGTTGGATTAAAGGAAGATGAGTAA 71980
QY 1000 1000
Db 71979 GAAATGGGGTGCAGGTGGAGGGGATAGGCATGACAGTCAAGAGTAACATGTAAGA 71920
QY 1000 1000
Db 71919 AATAATGATACTAGGACCATCTCTAAGGGGAAGTTAACACCTCTGTGTGGGATTTAC 71860
QY 1000 1000
Db 71859 GAGTAGACATCATTAATAAATGATTCAATAGAGCAGAGGTTTGAAAAGTCTTATC 71800
QY 1000 1000
Db 71799 ACAATGTCAGATCCATGTTGTACAGAACTCTTCAACCTTTTCCATGATTTTGGTCTT 71740
QY 1001 -----GlySerLeuThrGlnProLeuPheProCy 1010
Db 71739 ACTTTGGCCCATTTCTGCACCCCTCCACAGGAAGCCCTGACCCAGCCACTGTTTCCATG 71680
QY 1010 sLeuAlaAsnValSerAspProTyrLeuSerAlaValProGluGlySerAlaLeuIleCy 1030
|||||
Db 71679 CCTGCAATGATATCGACCATATTTGTCTGAGTCCAGAGGCGAGTGTGATCAG 71620
QY 1030 sLeu 1031
Db 71619 CTT-AGGTAAGCAGAATCAGTTTATAGAAACACTGCAACTCTAAAGAAATGTCACCTATTA 71561
QY 1032 -----GluAspLysArgLeuArgLeuValAspGlyAspSerArgCy 1045
Db 71560 ACTGACATTCTCCTCTAAGAGGACAAAGCGCTCCGCCCTAGTGGATGGGACAGCCGCTG 71501
QY 1045 sAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAspGlyTyr 1065
|||||
Db 71500 TGCCGGAGAGTAGATCTATCAGACGGCTTCTGGGCACTCTGTGATGACGGCTG 71441
QY 1065 pasLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAl 1085
|||||
Db 71440 GGACCTGAGCGATGCCACGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGCCCTTCAATGC 71381
QY 1085 aThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspLeuAsnCy 1105
Db 71380 CACGGTCTCTGCTCACTTTGGGGAGGGGTACAGGCCCACTGGCTGGATGACCTGAACCTG 71321
QY 1105 sThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCy 1125
|||||
Db 71320 CACAGGAATGGAGTCCCACTTGTGCGAGTGCCTTCCCGCGCTGGGGCAGCAGCTG 71261
QY 1125 sArgHisLysGluAspAlaGlyValIleCysSer 1136
|||||
Db 71260 CAGGCACAGGAGGAGCGAGGGGTCTCTGCTCAGGTGTGCTGCTACCTTTTCCACCT 71201
QY 1136 ----- 1136

Db 71200 CCTCTACACACACAGTGCACACACTTGCACACACACAACAATTTGTAAGTGTAGTGAAA 71141
QY 1136 ----- 1136
Db 71140 AGTGTATGGCCTTAGGAATACAAAATAATCCACTTGGTTTTTAATCCTTATTCTCTATA 71081
QY 1136 ----- 1136
Db 71080 CTGATTTCTTAATACCCCTTGTCTATATTACATTTGTCTGGGTATCATTCATAACTGAGAAG 71021
QY 1136 ----- 1136
Db 71020 AATAAAAGATTGGTCTTACTGTGAGAGATGTTAGACCCCAAGAAAGGACCTTACATCA 70961
QY 1136 ----- 1136
Db 70960 CATCOACTCTGGCTCTGAGAACACACTGTCAACCCGTGTCCAATCTGCCTTATTTTGTGT 70901
QY 1136 ----- 1136
Db 70900 TTTGTTTCTTTTATTTTAGGCTTTTAGTGCCTGAAGCCATGGTTATAGTTTCTGTCT 70841
QY 1136 ----- 1136
Db 70840 CTAGTGGTAAGAGAGAGAGGATGAGGAAGGTCTTTACTAACCTGACAGAGCAGAA 70781
QY 1136 ----- 1136
Db 70780 ACTAAGAACCCATAACTGTATCTCTCCCTTGGACACCCCTGCAANTGAGGTTAACCTCC 70721
QY 1136 ----- 1136
Db 70720 TGATTTATTTCTGAATTTGAATTTATGTTTCGAGAGTAGAATGTAGAAGGTGTCATTTAAG 70661
QY 1136 ----- 1136
Db 70660 TTAAGTTGATACAAATTTAAAGCTCAAATACTATTCTTTGGGAAACCCACACTCATCTGA 70601
QY 1136 ----- 1136
Db 70600 ANTCTACAAGACCCCTTAGTTTCCATCAGATTTATGCTGTGTTAATATGTTAGAGTGATTA 70541
QY 1136 ----- 1136
Db 70540 TGAATTTGTTGTGAGCCCTAATGTATTATTTCATGGGAAAAATGGATTCTCTTGGGA 70481
QY 1137 -----GluPh 1138
Db 70480 GGGATTTCCAACTCCAAGGAAATATTTAATACTAAGTTTGTATTCTCTGTAGAATT 70421
QY 1138 eThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluVa 1158
|||||
Db 70420 CACACCTTGGGCTCTACAGTGAACCTGAACACAGAGAGCTGTCTGGGAGATTGGAAGT 70361
QY 1158 lPheTyrAsnGlyThrTrpGlySerValGlyArgAsnIleThrThrAlaIleAlaGI 1178
|||||
Db 70360 CTTCTATAACGGACCTGGGCGAGCTGGCAGAGGAGCAATCACCACAGCCATAGCAGG 70301
QY 1178 yIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSe 1198
|||||
Db 70300 CATTGTGCGAGGAGCTGGGCTGTGGGAGAAATGAGTTGTCAGCCCTGCGCCCTTTATC 70241
QY 1198 rIystrThrGlySerGlyPheMetTrpValAspAspileGlnCysProLysThrHisIleSe 1218
|||||
Db 70240 TAAGACAGCTCTGGTTTCATGTGGTGGATGACATTCAGTCTCTCAAAACGCATATCTC 70181
QY 1218 rIleTrpGlnCysLeuSerAlaProTrpGluArgIleSerSerProAlaGluGluTh 1238
|||||
Db 70180 CATATGGCAGTGCCTGTCTGCCCATGGGAGCGAAGAAATCTCCAGCCCGCAGAGAGAC 70121
QY 1238 rTrpIleThrCysGlu 1243
|||||

Db 70120 CTGGATCACATGTGAAGTGGGTATCTTTCTAAACACAGCTGTCACTGAAGCCCTGTGA 70061
QY 1243 ----- 1243
Db 70060 AATATAACTTACCATAAATAGCAAAATCTTCCATTCAGGACACATGGCAGACTCACTGA 70001
QY 1243 ----- 1243
Db 70000 AAAGTGCATGACAATAAATGCTGATTTATCGCATCTGTCTATATGAAATTTTATTATT 69941
QY 1243 ----- 1243
Db 69940 TTGGTTTAGTATTTTACAGGGTATTTAAAAATGTATATGTCTCTTTTGATTCCTTTAAG 69881
QY 1243 ----- 1243
Db 69880 CTCCTTGTGAGAGCTGGATGAAGCACATANAAGGATAACACACATAAAAAATTACAG 69821
QY 1243 ----- 1243
Db 69820 TGTTTAAAAAATGAAAAATATAATATGTACTGTCTTAGAAATCTAGAAAGCTCAGA 69761
QY 1243 ----- 1243
Db 69760 ACTTTTGAGGAATGTGAATTAATTATGGAGATTAAAGCAGCTCTATTTTTCCTCGCAGAA 69701
QY 1243 ----- 1243
Db 69700 CATTTTACTATGGAAGTGTAGTTGATTATTTATAGACACATGTAAATGTCTTATTGA 69641
QY 1243 ----- 1243
Db 69640 TTAAAAATTAATATCTTTTATTCTAGAGATTTTAGAAATGTACAATTATAAATAATAAAA 69581
QY 1243 ----- 1243
Db 69580 TGCATTCCACGTTTAATCTGTATATCAAAAAAACCAATTACTPAATATTACGAAACATTAA 69521
QY 1243 ----- 1243
Db 69520 CTATTAGCTTCCTTTTGTACAAAAACTTAAAGTTATGATGTGTATATGAATGTGAT 69461
QY 1243 ----- 1243
Db 69460 ATGAAATATAGATACTATATATGTATATATATATAGTATGATATATATGATGATGAAT 69401
QY 1243 ----- 1243
Db 69400 TAGTGTACTACTATATTTAGGATATCTGTCTCTACTTTTATATGTAAAAAGTATACTGTA 69341
QY 1243 ----- 1243
Db 69340 AATATATTCTAATTTCAATGAATAGCCTTCAAAATGGCATACAAATATTTTATCATGTGAAC 69281
QY 1243 ----- 1243
Db 69280 ATAAATAACCATGTGATTTTGGTATTAGGCTATTTCTCAATGTCAAACTAGTATAAAT 69221
QY 1243 ----- 1243
Db 69220 AATGCCATTATAAGCATCCTTATTTCATTTTCGGTGAATCTCTGATATATATCTGAAGTTA 69161
QY 1243 ----- 1243
Db 69160 TATTTCTAGTCCCAAAATGAGTGGGTAAATGAGTGAACATTTTATGATCTTTATATG 69101
QY 1243 ----- 1243
Db 69100 TACTAGTTAAATGTCTCCAGCAGAAAGCCTATGTATTAGTTTTCATTTCTATCCCCAGTAT 69041
QY 1243 ----- 1243
Db 69040 ATGAGAGCATTGATTTTATTTCACCTTTCCCACTTCCCACTGTTACACACAGAGAAACATAT 68981

QY 1243 ----- 1243
Db 68980 ACACAGACATACACATACCTTCCCAAAATTAATAATTCACCTAAATCAACATAAATTTGCAT 68921
QY 1243 ----- 1243
Db 68920 SGGTTTGGTGTAGTAGTGATATGGACATATTTTATGAATATTTGTCTGTGTGGTTTATA 68861
QY 1243 ----- 1243
Db 68860 TTTTAAAGCATTTTCTCTCTTTTACTATTTGAATTTGTATCACAATAATATCATATCC 68801
QY 1243 ----- 1243
Db 68800 CAAACAAAAATTTGAAAAAACCCAGAAATCATGTGAAAAAACACAGATATCACCTATA 68741
QY 1243 ----- 1243
Db 68740 ATACCATGATCAGTTGATGATATATAGCTGATATTTACACTCTTGAGTATCATATGTGTG 68681
QY 1243 ----- 1243
Db 68680 TATATATTTATATATATTCATATATTTTATTTGCAAAATGTTATGCTTGCATATTTGTGC 68621
QY 1243 ----- 1243
Db 68620 TCCCTGCTCTTTTGATTAGCAATATTTTCTCAAGCTAAAGCATATATATTTTACCAT 68561
QY 1243 ----- 1243
Db 68560 ATGATATACTATAATCTTATGCACCAAAATCCCTGTCTTGGCCATCTAGAAAAATCAAGT 68501
QY 1243 ----- 1243
Db 68500 TATTTTTCATTTTAAAGTAACATCACATTAATAATCTTTAAATATATAGAAAAATATTTT 68441
QY 1243 ----- 1243
Db 68440 AAGCATATTTCTCAATATTTTCTCTCCAAAAATCTATTTTCTTATTTGATCAATATATTA 68381
QY 1243 ----- 1243
Db 68380 COTGAAATTTTGGATTTCAACTATCTTAAAGTTTAAATTTATCAATATTTTATATAAAAA 68321
QY 1243 ----- 1243
Db 68320 TTTGCACTATATATGATATGAATGATACATATGCAGTAGATCTTATTTAGCGCTCTTTTA 68261
QY 1243 ----- 1243
Db 68260 TATTTTGTGATATTAGCATTTGGAGATTTCTGAAATATAAACTAGATGAATTCGCCAG 68201
QY 1243 ----- 1243
Db 68200 CTTTATGATCTTTAATAGTATATGAAATAAAATCATTTGTGTTTCAATATTTGAAAGAGA 68141
QY 1243 ----- 1243
Db 68140 ATAACTATAGCTATTTTGGGTAGTAAATTTGTTAAATCTTAAAAAGAAATGTCAGGACTAT 68081
QY 1243 ----- 1243
Db 68080 TGGCTTATTTCTCATTTTCTTGAGTCAATTTTGTGAAGGATGAAAGATGTATTAGTCAG 68021
QY 1243 ----- 1243
Db 68020 TTCTCATGCTGCTAATAAAGAAATACCAGACTGGGTAATTTTATAAAGGAAGAGGCTT 67961
QY 1243 ----- 1243
Db 67960 NATGGACTCACAGTTCCCATATGGCTTAATGAGGCCTCACAATCATGGCTGAAGGCAAGGA 67901


```
Db 65740 TGAATCACTGAATGCCTCCTCAGGTATAGTCTTCTGCTATCAGTGGCGACACAAATAC 65681
QY 1357 ----- 1357
Db 65680 CTAAGGCTACATGGAAGGTCAAGGGAGGCCAGAACTGCTATGATAATCAAGAAGACTG 65621
QY 1357 ----- 1357
Db 65620 GGGACTGCAGAAATATAGTCTGTGGTGGCGAGCCAGTGGTATCAATAGCTCTGTGACTG 65561
QY 1357 ----- 1357
Db 65560 CTTTGTATGCAGTCTTCTCTCTAGAAAAATCCTTTTACATGCAGAGAGGATAAGAC 65501
QY 1357 ----- 1357
Db 65500 TCTCATGTGACCTCATAGATAATCTTCTAGATCTCAAAATATCTTTATGAGACTGCG 65441
QY 1357 ----- 1357
Db 65440 AGAAACAGCTCAAGCTTCAGGAGAAATCACTTAAACCCAGGAAGCAGAGATTGCACTGAGC 65381
QY 1357 ----- 1357
Db 65380 CGAGATCGTCCAGTACACTCAGCCTGGGTGACAGAAATGAGACTCCATCTCAAAAAAAA 65321
QY 1357 ----- 1357
Db 65320 AAAAAAGAAAGAACAGCTCAGCTCAAGCCCTAGTATTTCTGATTCGTGGTGTCTTAGT 65261
QY 1357 ----- 1357
Db 65260 AGTGAATGCAATGCAAGAGAAAGGTTAGTTATCAAAATCTCTGTATAAAATTTGAAGT 65201
QY 1357 ----- 1357
Db 65200 AATTTAGGCAATAGAAATATGGAAGATTTGGATTCTGTTCCAGATAGTTTTTCAT 65141
QY 1357 ----- 1357
Db 65140 CTAAATAGCTCTCAAGATTAAGTACTGTTCTCATGAATATGTTGAATGTATAGCAATTA 65081
QY 1357 ----- 1357
Db 65080 TTCTTCTCAGCAACTAGGAACACGTTTCTTTTACAATGTATAAACCCACAATTTGATATC 65021
QY 1358 ----- GlyHisLeuAlaLeuIleLeus 1365
Db 65020 TACTAGTTAATATCAATGCTTAATTTCTATACAGGTCATTTAGCACTTATTTTAT 64961
QY 1365 erSerIlePheGlyLeuLeuValLeuPheIleLeuPheLeuThrTrpCysArgV 1385
Db 64960 CCAGTATCTTTGGGCTCCTTCTCTGGTCTCTGTTATCTTATTTCTCAGTGTGTCGCCAG 64901
QY 1385 aLgInLySgInLySHisLeuProLeuArgVal----- 1395
Db 64900 TTCAGAAAACAAAACATCTGCCCTCAGAGG-TGAGAGCCCACTAGGATTGTATAAAAT 64842
QY 1395 ----- 1395
Db 64841 GTGAGTTACCTGCCACCAAGGCCACAGAGTGGGGATCTTTGAGAAATGGAAGAAATG 64782
QY 1395 ----- 1395
Db 64781 CACTTGCCACTTTTGTGTGATGATGGATTGGAAGTCCAGGCCCAATGCAGAAACATGTGG 64722
QY 1395 ----- 1395
Db 64721 AATGATCAAGAAATCAGCCAATCCACAATTTCCATTTACACAAATGTAATCTTTTAGAGC 64662
QY 1395 ----- 1395
Db 64661 TGTAGCTTGCCATACAGACGATGAGTCAGATAGGCAAAATATCTTAGCTGATAGAAGGAG 64602
```

```
QY 1395 ----- 1395
Db 64601 GTAATTTGTATGTTTGGTTCAGAAATCTAGACTCCAGAAATAGCCATAGAAAATGTAGAA 64542
QY 1395 ----- 1395
Db 64541 GCATGCTGGAGGATTAAAGAACAAGAGTTTCCCTGGGTTTGAACACCACCTTATATAAATTAC 64482
QY 1395 ----- 1395
Db 64481 TAGCTGTGTGGCCTTGGCTAGAAACTACCTTTTCCAAACTCCAAATTTCCCTAATTTGTAAA 64422
QY 1395 ----- 1395
Db 64421 GATAATAATAGTCTTTACCTCATCAGGTTAAGATAGTTAAATAAATTAATCATCTGTATGA 64362
QY 1395 ----- 1395
Db 64361 AAAATGGCCAGCACATAGTAAATACTATGTAAAGAACTTTGGTAAATAAAGTGAAGAATA 64302
QY 1395 ----- 1395
Db 64301 TCAAGGGAGAAAAGAGGAGGATAAATTTTAAAAATGTAAAAAATTCACTAGCAGGGGTAGG 64242
QY 1395 ----- 1395
Db 64241 TGCTGAAAGAGCCATTGGATTTTAGCTCAGAGGGATTAAATATGTGTCAAAACCTTAGCAAT 64182
QY 1395 ----- 1395
Db 64181 AAGTGTCTTAATATAAGGGATAGGCCATTTGAATTTTGAATCCTTAACCATGAAGCATCCAT 64122
QY 1396 ----- SerThrArgArgArgGlySerLeuGluIuAsnLeuPheHisGluMetGluT 1413
Db 64121 TCTCAGTTTCAACCAAGAGAGGGGTTCTCTCGAGGAGAAATTTATTTCCATGAGATGGAGA 64062
QY 1413 hrCysLeuLysArgGluAspProHisGlyThrArgThrSer 1426
Db 64061 CCTGCCTCAAGAGAGAGGAGGCCACATGGGACAAAGAACCTCA 64021
```

Search completed: May 12, 2003, 19:20:08
Job time : 6250 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 06:24:52 ; Search time 52 Seconds
(without alignments)
5757.432 Million cell updates/sec

Title: US-09-759-130B-381

Perfect score: 8138

Sequence: 1 MMLPQNSWHIDFGRCCHQON.....CEDASDTSLGLVPASEATK 1453

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriopl:*
- 17: sp_cheap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8109	99.6	1453	Q9NR16	Q9nr16 homo sapien
2	3441.5	42.3	1151	Q07899	Q07899 homo sapien
3	3429	42.1	1116	Q07898	Q07898 homo sapien
4	3429	42.1	1156	Q07900	Q07900 homo sapien
5	3396.5	41.7	1149	Q07901	Q07901 homo sapien
6	3272	40.2	1121	Q99MX8	Q99mx8 mus musculu
7	2990	36.7	2403	Q9UGM2	Q9ugm2 homo sapien
8	2987.5	36.7	2426	Q9UGM3	Q9ugm3 homo sapien
9	2987	36.7	2412	Q9UJ57	Q9uj57 homo sapien
10	2982	36.6	2413	Q96DU4	Q96du4 homo sapien
11	2980	36.6	2413	Q9UKJ4	Q9ukj4 homo sapien
12	2448	30.1	1785	Q9Y211	Q9y211 homo sapien
13	2439.5	30.0	2153	Q9Y375	Q9y375 strongyloce
14	2435	29.9	1785	Q9Y4V9	Q9y4v9 homo sapien
15	2246.5	27.6	2043	Q96943	Q96943 geodda cydo
16	2156.5	26.5	1594	Q95218	Q95218 oryctolagus

17	1962.5	24.1	2083	11	Q60997	Q60997 mus musculu
18	1845	22.7	1957	11	Q9JMJ9	Q9jnj9 mus musculu
19	1594.5	19.6	1188	5	Q9NC87	Q9nc87 strongyloce
20	1586	19.5	710	6	Q95JES	Q95jes ovis aries
21	1431.5	17.6	749	5	Q9NC88	Q9nc88 strongyloce
22	1399	17.2	804	6	Q29113	Q29113 sus scrofa
23	1341	16.5	546	6	Q29111	Q29111 sus scrofa
24	1301	16.0	1075	5	Q9NC90	Q9nc90 strongyloce
25	1277	15.7	973	5	Q9NC89	Q9nc89 strongyloce
26	1219	15.0	369	6	Q29110	Q29110 sus scrofa
27	1205	14.8	420	6	Q95JD9	Q95jd9 ovis aries
28	1085.5	13.3	480	6	Q29112	Q29112 sus scrofa
29	1078.5	13.3	600	6	Q28910	Q28910 bos taurus
30	1048.5	12.9	1290	11	Q62827	Q62827 rattus norv
31	1041.5	12.8	470	13	Q98TD1	Q98td1 gallus gall
32	968.5	11.9	575	4	Q8WTU2	Q8wtu2 homo sapien
33	866.5	10.6	1036	5	Q97378	Q97378 strongyloce
34	763.5	9.4	437	6	Q29109	Q29109 sus scrofa
35	723.5	8.9	749	11	Q99JC8	Q99jc8 rattus norv
36	703	8.6	209	6	Q28881	Q28881 ovis aries
37	686.5	8.4	457	5	Q9NC91	Q9nc91 strongyloce
38	665	8.2	918	13	Q92098	Q92098 petromyzon
39	664.5	8.2	268	6	Q9TV20	Q9tv20 oryctolagus
40	664.5	8.2	277	6	Q9TV21	Q9tv21 oryctolagus
41	662.5	8.1	264	6	Q9TV22	Q9tv22 oryctolagus
42	650	8.0	505	4	Q96180	Q96180 homo sapien
43	632	7.8	754	11	Q91VN8	Q91vn8 mus musculu
44	620.5	7.6	665	11	Q91WN5	Q91wn5 mus musculu
45	603.5	7.4	592	4	Q8WWJ7	Q8wwj7 homo sapien

ALIGNMENTS

RESULT 1

Q9NR16	ID	Q9NR16	PRELIMINARY;	PRT;	1453 AA.
AC	Q9NR16;				
DT	01-OCT-2000 (TReMBLrel. 15, Created)				
DT	01-OCT-2000 (TReMBLrel. 15, Last sequence update)				
DT	01-MAR-2002 (TReMBLrel. 20, Last annotation update)				
DE	Scavenger receptor cysteine-rich type 1 protein M160 precursor.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-20540099; PubMed-11086079;				
RA	Gronlund J., Vitved L., Lausen M., Skjodt K., Holmskov U.;				
RT	"Cloning of a Novel Scavenger Receptor Cysteine-Rich Type I				
RT	Transmembrane Molecule (M160) Expressed by Human Macrophages."				
RL	J. Immunol. 165:6406-6415(2000).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Gronlund J., Vitved L., Lausen M., Skjodt K., Holmskov U.;				
RL	Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.				
DR	EMBL: AF264014; AAF91396.1; -				
DR	InterPro: IPR001190; Srcr_receptor.				
DR	Pfam: PF00530; SRCR_12.				
DR	PRINTS; PR00258; SPERACTRCPTR.				
DR	SMART; SM00202; SR; 11.				
DR	PROSITE; PS00420; SRCR_1; UNKNOWN_5.				
DR	PROSITE; PS50287; SRCR_2; 12.				
KW	Receptor; Signal.				
FT	SIGNAL 1 42				POTENTIAL.
SQ	SEQUENCE 1453 AA; 159263 MW; 15445CE9868CF58A CRC64;				

Query Match 99.6%; Score 8109; DB 4; Length 1453;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 NMLPQNSWHIDFGRCCHQONLFSAVVTCTILLNSCFLISSFNGTDLRLVNGDPCSGT 60

```
Db 1 MMLPQNSWHDGRCCCHONLFSAVVTCILLNSFLISSFNGTDLRLVNGDPCSGT 60
QY 61 VEYKFGQGTGTCDDGWNNTASTVYVCKQLGCPFSFAMFRFGQAVTRHGKIWLDDVSCYGN 120
Db 61 VEYKFGQGTGTCDDGWNNTASTVYVCKQLGCPFSFAMFRFGQAVTRHGKIWLDDVSCYGN 120
QY 121 ESALWPCQHRHNGSHNCYHGEDVGNVCYGEANLGLRLVDGNNSCGRVEVKFOERWGTIC 180
Db 121 ESALWPCQHRHNGSHNCYHGEDVGNVCYGEANLGLRLVDGNNSCGRVEVKFOERWGTIC 180
QY 181 DQGNLNTAAVVCROLGCPSSFISSGVNNSPAVLRPIWLDDILCOGNELALNCHRHGNG 240
Db 181 DQGNLNTAAVVCROLGCPSSFISSGVNNSPAVLRPIWLDDILCOGNELALNCHRHGNG 240
QY 241 NHDCSHNEDVTLCYDSSDLRLVGGTNRCMGRVELKIQGRWGTVCVCHKWNNAADVVC 300
Db 241 NHDCSHNEDVTLCYDSSDLRLVGGTNRCMGRVELKIQGRWGTVCVCHKWNNAADVVC 300
QY 301 KQLGCGTALHFAGLPHLOSGSDVWLDGVSCGNSFVWDCRHSGTVPNFDCLHQNDVSI 360
Db 301 KQLGCGTALHFAGLPHLOSGSDVWLDGVSCGNSFVWDCRHSGTVPNFDCLHQNDVSI 360
QY 361 CSDDGADLELRADGNNCSGRVEVRIHQWWTICDQNKNEQALVYCKQLGCPFSFVGSR 420
Db 361 CSDDGADLELRADGNNCSGRVEVRIHQWWTICDQNKNEQALVYCKQLGCPFSFVGSR 420
QY 421 RAKPSNEARDIWNISCTGNSALWDCITYDGKAKRTCFRRSDAGVICSADKADLRLVG 480
Db 421 RAKPSNEARDIWNISCTGNSALWDCITYDGKAKRTCFRRSDAGVICSADKADLRLVG 480
QY 481 AHSPCYGRLEVYKQGEWGTVCVCHDRSTRNAAVYCKQLGCGKPMHVFGMTYFKEASGPIWL 540
Db 481 AHSPCYGRLEVYKQGEWGTVCVCHDRSTRNAAVYCKQLGCGKPMHVFGMTYFKEASGPIWL 540
QY 541 DVVSCIGNESNTWDCSHSGWKHNCVHREDVITVCSGATWGLRLVGSNCRSCGRLEVYF 600
Db 541 DVVSCIGNESNTWDCSHSGWKHNCVHREDVITVCSGATWGLRLVGSNCRSCGRLEVYF 600
QY 601 QGRWGTVCDDGWNKAAAVVCSQDPCSSIIGMGLGNASTGYGKIWLDDVSCDGEDSLW 660
Db 601 QGRWGTVCDDGWNKAAAVVCSQDPCSSIIGMGLGNASTGYGKIWLDDVSCDGEDSLW 660
QY 661 SCRNCSGWNNDCHSHEDVGVICSDASDMLRLVGGSSRCAGKVEVYVNOGAVGILCANGW 720
Db 661 SCRNCSGWNNDCHSHEDVGVICSDASDMLRLVGGSSRCAGKVEVYVNOGAVGILCANGW 720
QY 721 MNIAEVVCRQLCGSAIRVSRPHEFTERTLHILMSNSGCTGGEASLWDCIKWENKOTACH 780
Db 721 MNIAEVVCRQLCGSAIRVSRPHEFTERTLHILMSNSGCTGGEASLWDCIKWENKOTACH 780
QY 781 LNWEASLICSARHQRPLVGDMPSCSGRVEVKHADTWRSVCDSDFSLSHAANVLCRELNGD 840
Db 781 LNWEASLICSARHQRPLVGDMPSCSGRVEVKHADTWRSVCDSDFSLSHAANVLCRELNGD 840
QY 841 ATLSLVGDFHFGKNGLTWAEKFCQSGSETHLALCPVIOHPEDTCHSREVGVVCSRYTIV 900
Db 841 ATLSLVGDFHFGKNGLTWAEKFCQSGSETHLALCPVIOHPEDTCHSREVGVVCSRYTIV 900
QY 901 RLNVNKSQCDGVEINVLGHWSGLCDTHWDPEDARVLCRLSCGTALSTTGKYGIGERSV 960
Db 901 RLNVNKSQCDGVEINVLGHWSGLCDTHWDPEDARVLCRLSCGTALSTTGKYGIGERSV 960
QY 961 RVWGHFRHCLGNESLLDNCQMTVLGAPPCIRHNTVSVICTGSLTQPLFPCLANVSDPYLS 1020
Db 961 RVWGHFRHCLGNESLLDNCQMTVLGAPPCIRHNTVSVICTGSLTQPLFPCLANVSDPYLS 1020
QY 1021 AYPEGSAIICLEDKRLRLVGDGSRACAGVEIYHDFGWTICDDGWDLSDAHVVCKLCCG 1080
Db 1021 AYPEGSAIICLEDKRLRLVGDGSRACAGVEIYHDFGWTICDDGWDLSDAHVVCKLCCG 1080
QY 1081 VAFNATVSAHFEGSGPIWLDDNCTGTESHLWQCPSGRWGQDCHRHKEDAGVICSEFTA 1140
Db 1081 VAFNATVSAHFEGSGPIWLDDNCTGTESHLWQCPSGRWGQDCHRHKEDAGVICSEFTA 1140
```

```
Db 1081 VAFNATVSAHFEGSGPIWLDDNCTGMESHLMWQCPSGRWGQDCHRHKEDAGVICSEFTA 1140
QY 1141 LRLYSETETESCAGLEVFYNGTWSVGRNRNITTAIAGIVCRQLCGGNGVYSLAPLSKT 1200
Db 1141 LRLYSETETESCAGLEVFYNGTWSVGRNRNITTAIAGIVCRQLCGGNGVYSLAPLSKT 1200
QY 1201 GSGFMVDDIOCPKTHISIWQCLSAWERRISSPAEETWITCEDRIRVRGDPTECSGRVE 1260
Db 1201 GSGFMVDDIOCPKTHISIWQCLSAWERRISSPAEETWITCEDRIRVRGDPTECSGRVE 1260
QY 1261 IWHAGSWGTVCCDDNDLAEAEVVCQOLGCGSALAALRDASFOQGTGTIWLDDMRCKGNES 1320
Db 1261 IWHAGSWGTVCCDDNDLAEAEVVCQOLGCGSALAALRDASFOQGTGTIWLDDMRCKGNES 1320
QY 1321 FLWDCCHAKPWGSDCGHKGEDAGVRCGSLKSLNASSGHLALILSSIFGLLLVLFILFL 1380
Db 1321 FLWDCCHAKPWGSDCGHKGEDAGVRCGSLKSLNASSGHLALILSSIFGLLLVLFILFL 1380
QY 1381 TWCRVQKQKHLPLRVSTRRRGSLLENLFHEMETCLKREDPHGTRTSDDTPNHGCEADST 1440
Db 1381 TWCRVQKQKHLPLRVSTRRRGSLLENLFHEMETCLKREDPHGTRTSDDTPNHGCEADST 1440
QY 1441 SLLGVLPASEATK 1453
Db 1441 SLLGVLPASEATK 1453

RESULT 2
Q07899
ID Q07899 PRELIMINARY; PRT: 1151 AA.
AC Q07899;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE M130 antigen cytoplasmic variant 1 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93380506; PubMed=8370408;
RA Law S.A., Micklem K.J., Shaw J.M., Zhang X.P., Dong Y., Willis A.C.,
RA Mason D.Y.;
RT "A new macrophage differentiation antigen which is a member of the
RT scavenger receptor superfamily."
RL Eur. J. Immunol. 23:2320-2325(1993).
DR EMBL: 222969; CAAB0542.1; -
DR InterPro: IPR001190; Scrv_receptor.
DR Pfam: PF00530; SRCR; 9.
DR PRINTS: PR00258; SPERACTRCPTR.
DR SMART: SM00202; SR; 9.
DR PROSITE: PS00420; SRCR_1; 4.
DR PROSITE: PS50287; SRCR_2; 9.
KW Signal.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 1151 AA; 124820 MW; 3CDBAD857B6EB91 CRG64;

Query Match 42.3%; Score 3441.5; DB 4; Length 1151;
Best Local Similarity 55.2%; Pred. No. 4e-260;
Matches 608; Conservative 166; Mismatches 311; Indels 17; Gaps 6;

QY 364 GADLELRADGNNCSGRVEVRIHQWWTICDQNKNEQALVYCKQLGCPFSFVGSRRAK 423
Db 41 GTDELRLVDGNNCSGRVEVVKVQEWGTVCNNGWSMEAVSVICNQLGCPAIKAPGAN 100
QY 424 PSNEARDIWNISCTGNSALWDCITYDGKAKRT-CFRRSDAGVICSADKADLRLVGAH 482
Db 101 SSAGSGRIWMDHVSCKGNSALWDCIKHDGKWSNCTHQDAGVTCSDGNSLMRLTRGG 160
QY 483 SPCYGRLEVYKQGEWGTVCVCHDRSTRNAAVYCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
Db 483 SPCYGRLEVYKQGEWGTVCVCHDRSTRNAAVYCKQLGCGKPMHVFGMTYFKEASGPIWLDD 542
```



```
Db 161 NMSGRIEIKFQGRWGTCCDNFNIDHASVTCROLECGSAVSFGSSNFGGSGPIWFD 220
Qy 543 VSCIGNESNIWDCSHSGWGHNCVHREDVIVTCSDATWGLRLVGGNRCGRLEVFQ 602
Db 221 LICNGNESALWNCKHOGKHNCDHAEDAGVICSKGADLSRLVDGYTECSGRLEVFQ 280
Qy 603 RWGTVCDGWNKSKAAAVVCSOLDPCSSITIGMGLNASTGYGKIWLDDVSCDGDSDLSWC 662
Db 281 EWGITCDGWDSDYDAAVACKQLGCPATAVTAIGRVNASKGFGHIWLDVSQCGHEPAWQC 340
Qy 663 RNSGWNDCSHSDVGVICSDASDMELRLVGGSSRCAGKVEVNVQAVGILLCANGCMN 722
Db 341 KHEWGHYCNHNDAGVTCSDGDLRLURGGSGRCAGTVEIQRLGKVCDRGWLK 400
Qy 723 IAEVVCROLECGSAIRVSREPHFTERTLHILMSNGTGGGEASLWDCIRWEWKOTACHLN 782
Db 401 EADVVCRLGCGSALKTSYQVYSKIQTNTWLFSSCNGNETSLWCKNQMGGTLCDHY 460
Qy 783 MEASLICASHPRLVADPCSGRVEVKADTWRSVCDSDFSLSHAANVLCRELNCGD 842
Db 461 EAKITCSAHPRLVADPCSGRVEVKADTWRSVCDSDFSLSHAANVLCRELNCGD 520
Qy 843 SLVGDHFGKGLUTWAEKFCGSETHLALCPVQHPEDTCIHSREVVGVCSRYTDVRL 902
Db 521 SILGAHFGGNGGQIWAEEFQCEGHESHLSCVPAPREGTCSHSRDVGVVCSRYTEIRL 580
Qy 903 VNGKSQDGOVVEINVLGHWSLCTHWDPEARVLCRLQSGTALSTTGGKYIGERSVRV 962
Db 581 VNGKTPCEGRVELKTLGWSLCSNHWIDEDAHVLCQOLKCGVALSTPGGARFCGNGOI 640
Qy 963 WGRHFCGLNESLLDNCOMTVLGAPPCHGNTVSVITGTSUPLFPCLANVSDPYLSAV 1022
Db 641 WRHMFHCTGTEQHMGCDPVATGALGASLPCSEOVASVICSGNQSQTLSNCSLGPTRPTI 700
Qy 1023 PEGSALICLEDKRLRLVGDGSRACGRVEIYHGWFGWTCDDGWDLSDAHVVVCRLGCGVA 1082
Db 701 PEESAVACIESGQRLRLVNGGRCAGRVEIYHGSWGTICDDSDWLSDAHVVVCRLGCGVA 760
Qy 1083 FNATVSAHFEGGSGPIWLDLNCITGTESHLWQCPSRGWQHDCRKHEDAGVICSEFTALR 1142
Db 761 INATGSAHFEGGSGPIWLDLNCITGTESHLWQCPSRGWQHDCRKHEDAGVICSEFTALR 820
Qy 1143 LYSTETESACGRLEVFYNGTWGVSGRNNTTATAGIVCRLQCGGNGVWSLAPLSKTGS 1202
Db 821 LTSEASREACAGRLEVFYNGTWGVSGRNNTTATAGIVCRLQCGGNGVWSLAPLSKTGS 880
Qy 1203 GFMAVDIOQCPKTHISTWQCLSAPEWERRISSPAETWITCEDRIRVGGDTCEGSRVEIW 1262
Db 881 IPMVDNVQCPKGPDTLWQCPSPSWERKLASPSSETWITCNKRLQEGPTSCSGRVEIW 940
Qy 1263 HAGSWGTVCDSDWDLAEAVVVCQQLGCGSALALRADSFQGGTGIWLDMDRCKGNESFL 1322
Db 941 HGSWGTVCDSDWDLDAQVVCQQLGCGPALKAFKEAFEGGDTGPIWLNVEVKCKGNESL 1000
Qy 1323 WDCHAKPWGSDCHGKEDAGVRCSSQSLK-----SLNASSGHALILSSIFGLLLVLF 1376
Db 1001 WDCPARWGHSECHKEDEAVNCTDISVQKTPQKATTGRSSRSFSAVGLGLVLLAIF 1060
Qy 1377 I--LFLTWCRVOKHPLRVTRRRGSLLENLPHMETCLKREDPHGTRTSDTTPNHGC 1434
Db 1061 VALEFLTKRRQROR---LAVSSRGENLVHQIYQIRREMNSCLNADDLMLNSSEN--SHES 1115
Qy 1435 EDASDTSLGV--LPASEATK 1453
Db 1116 ADFSAAELISVKFLPISGMEK 1137
```

RESULT 3

Q07898

ID Q07898 PRELIMINARY; PRT: 1116 AA.

AC Q07898;

DT 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)

```
DT 01-MAF-2002 (Tremblrel. 20, Last annotation update)
DE M130 antigen precursor.
GN CD163.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
...OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93380506; Pubmed=8370408;
RA Law S.A., Micklem K.J., Shaw J.M., Zhang X.P., Dong Y., Willis A.C.,
RA Mason D.T.;
RT "A new macrophage differentiation antigen which is a member of the
RT scavenger receptor superfamily."
RL Eur. J. Immunol. 23:2320-2325(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Ritter M., Buehler C., Langmann T., Schmitz G.;
RT "Genomic organization of the human CD163 gene."
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z22968; CAA80541.1;
DR EMBL; Y18388; CAB45233.1;
DR EMBL; Y18389; CAB45233.1; JOINED.
DR EMBL; Y18390; CAB45233.1; JOINED.
DR EMBL; Y18391; CAB45233.1; JOINED.
DR EMBL; Y18392; CAB45233.1; JOINED.
DR EMBL; Y18393; CAB45233.1; JOINED.
DR EMBL; Y18394; CAB45233.1; JOINED.
DR EMBL; Y18395; CAB45233.1; JOINED.
DR EMBL; Y18396; CAB45233.1; JOINED.
DR EMBL; Y18397; CAB45233.1; JOINED.
DR EMBL; Y18398; CAB45233.1; JOINED.
DR EMBL; Y18399; CAB45233.1; JOINED.
DR EMBL; Y18400; CAB45233.1; JOINED.
DR EMBL; Y18401; CAB45233.1; JOINED.
DR EMBL; Y18402; CAB45233.1; JOINED.
DR EMBL; Y18403; CAB45233.1; JOINED.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00530; Srcr; 9.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00202; SR; 9.
DR PROSITE; PS00420; SRCR_1; 4.
DR PROSITE; PS00287; SRCR_2; 9.
KW Antigen; Signal.
FT SIGNAL 1 40 POTENTIAL.
FT CHAIN 1 1116 M130 ANTIGEN.
SQ SEQUENCE 1116 AA; 120979 MW; E805BC80B52E4009 CRC64;

Query Match 42.18; Score 3429; DB 4; Length 1116;
Best Local Similarity 56.28; Pred. No. 3.7e-259;
Matches 598; Conservative 161; Mismatches 294; Indels 12; Gaps 4;

Qy 364 GADLELRADGNNCSGRVEVRIHEOWMTICDQWKNQALVVCVKOLGCPFSVFGSRRAK 423
Db 41 GTDKELRLVDGKNCGRVEVVKVQEEWGTVCNNGMSAEVSVICNLQCPCTAIRAPGAN 100
Qy 424 PSNEARDIWNISICTGNESALWDCTYDGAART--CFRRSDAGVICSADLRLVGAH 482
Db 101 SSAGSGRLWMDHVSCRGNESALWDCKHDGKHSNCTHQDAGVTCSDGSLNEMRLTRGG 160
Qy 483 SPYGRLEVKYQGEWGTVCHDRWSTRNAVVCVKOLGCKGPMHVFQMTYFKBASGPIWLD 542
Db 161 NMSGRIEIKFQGRWGTCCDNFNIDHASVTCROLECGSAVSFGSSNFGGSGPIWFD 220
Qy 543 VSCIGNESNIWDCSHSGWGHNCVHREDVIVTCSDATWGLRLVGGNRCGRLEVFQ 602
Db 221 LICNGNESALWNCKHOGKHNCDHAEDAGVICSKGADLSRLVDGYTECSGRLEVFQ 280
Qy 603 RWGTVCDGWNKSKAAAVVCSOLDPCSSITIGMGLNASTGYGKIWLDDVSCDGDSDLSWC 662
Db 281 EWGITCDGWDSDYDAAVACKQLGCPATAVTAIGRVNASKGFGHIWLDVSQCGHEPAWQC 340
Qy 663 RNSGWNDCSHSDVGVICSDASDMELRLVGGSSRCAGKVEVNVQAVGILLCANGCMN 722
```

```
Db 341 KHEWGHKYNHEDAGVTCSDGSDLELRKRGSGRCAGTVEVEIQRLRLGKVCDDRGWGLK 400
Qy 723 LAEYVCRQLECGSALRVSRPHFTERTLHILMSNGCTGGASLWDCILRWKWKQTACHLN 782
Db 401 EADVCRQLGCGSALKTSYQVYSKIQATNTWLFSSCNGNETSLWDCNKGWGLTCDHY 460
Qy 783 MEALICSARHQPRLVGADMPGSRVYKVDHATWRSVCDSDFSLHAANVLCRELNCDAI 842
Db 461 BEAKITCSAHPRELVGGDIPCSGRVYKVDHATWRSVCDSDFSLEAASVLCRELQCGTW 520
Qy 843 SILVGDHFGKNGLTWAEKFCOEGSETHALCPVQHEDTCIHSREVGVCSRYTDVRL 902
Db 521 SILGAHFGEGNGQIWAEEFCOEGSETHALCPVAPRPEGTCSHSRDVGVCSTYETRL 580
Qy 903 VNGKSCDQGVQVEINVLGHWSGLCDTHWDPEDARVLCRLSCGTALSTGGKYIGERSVRV 962
Db 581 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKNGQI 640
Qy 963 WGRPHCLGNESLNDNCOMTVLGAPPCIHGNTVSVICTGSLTQPLPCLANVSDPYLSAV 1022
Db 641 WRHMFHCTGTQEHMGDCPVTALGASLCPSEQVAVSICSGNSQTLSSCNSSSLGTRPTI 700
Qy 1023 PEGSALICLEDKRLRLVGDGRSACGRVEIYHDGFWGTTICDDGWLSDAHVVCQKLGCGVA 1082
Db 701 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGWSGTTICDDSWLSDAHVVCRLGCGEA 760
Qy 1083 FNATVSAHFEGSGPILWDLNCTGESHLMQCPGSRGQHDCHKEHDAGVICSFTALR 1142
Db 761 INATGSAHFEGSGPILWDLNCTGESHLMQCPGSRGQHDCHKEHDAGVICSFTALR 820
Qy 1143 LYSETETESACRLEVFYNGTWGSGVRNITTAIAGIVCRLGCGENGVSAPLSKGTGS 1202
Db 821 LTSEASREACAGRLVFNAGWTVGKSSMSETTVGVVCRQLGCAKDKINPASILDKAMS 880
Qy 1203 GFWMVDDIQCPTKTHISIWQCLSAWERRISSPAETWTTCEDRIVRGDGTCSGRVFIW 1262
Db 881 IPMWVDDIQCPTKTHISIWQCLSAWERRISSPAETWTTCEDRIVRGDGTCSGRVFIW 940
Qy 1263 HAGSGTVCDDSDWDLAEAVVCCQLGCGSALAAALDASFGQGTGTIWLDDMRCKGNESFL 1322
Db 941 HGGSGTVCDDSDWDLDDAQQVCCQLGCGPALKAPEAFEGGTGPIWLVNCKGNESFL 1000
Qy 1323 WDCHAKPGQSDCHKEHDAGVICSQSLK-----SLNASSGHIALIUSIFGLLLVLF 1376
Db 1001 WDCPARRGHSCECHKEHDAVNACTDISVQKTPQKATTCGRSSRQSSFIAGVILGVVLLAIF 1060
Qy 1377 I--LFLTWCRVQKOKHPLRVSTRRRGSLLENLPHMETCLKRED 1419
Db 1061 VALFELTKRRQRQR---LAVSSRGENVLHQIYQRMNSCLNADD 1102
```

RESULT 4

```
Q07900 PRELIMINARY; PRP; 1156 AA.
AC Q07900;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE M130 antigen cytoplasmic variant 2 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RX [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93380506; PubMed=8370408;
RA Law S.A., Micklem K.J., Shaw J.M., Zhang X.P., Dong Y., Willis A.C.,
RA Mason D.Y.;
RT "A new macrophage differentiation antigen which is a member of the
RL scavenger receptor superfamily.";
RL Eur. J. Immunol. 23:2320-2325(1993).
DR EMBL; Z22970; CAA80543.1; -
```

```
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00530; SRCR; 9.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00202; SR; 9.
DR PROSITE; PS00420; SRCR_1; 4.
DR PROSITE; PS50287; SRCR_2; 9.
KW Signal.
FT CHAIN 1 40 POTENTIAL.
FT CHAIN 41 1156 M130 ANTIGEN CYTOPLASMIC VARIANT 2.
SQ SEQUENCE 1156 AA; 125352 MW; 3BEEC22D805BD5C CRC64;
```

Query Match 42.1%; Score 3429; DB 4; Length 1156;
Best Local Similarity 56.2%; Pred. No. 3.8e-259;
Matches 598; Conservative 161; Mismatches 294; Indels 12; Gaps 4;

```
Qy 364 GADLELRADGNSNCGSRVEVRIHEQWWTICDNNKNQALVVCQKLGCPFSVFGSRRAK 423
Db 41 GTDKELRLVGDGRSACGRVEIYHDGFWGTTICDDGWLSDAHVVCQKLGCGVA 100
Qy 424 PSNEARDIWINISICTGNESALWDCTYDGKAKRT-CFRRSDAGVICSADKADLDRLVGAH 482
Db 101 SSAGSGRIWMDHVSCRGNESALWDCKHDGKHSNCTHQQDAGVTCSDGSLNEMRLTRGG 160
Qy 483 SPCYGRLEVYKQEGMTVCHDRMSTRNAAVVCQKLGCGKPMHVFQMTYFKEASGPIWLDD 542
Db 161 NMCSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLECGSAVSFGSSNFGEGSGPIWDD 220
Qy 543 VSCIGNESINWDCSHGKNGKUNCHVREDIVVTCSDATWGLRLVGSNRCGRLEVYFQG 602
Db 221 LICNGESALWCKHQGWKHNCHDAEDAGVICSAGDLSURLVDVGTCTCSGRLEVRFG 280
Qy 603 RWGTVCDDGWSKAAAVVCSQDLCPSSTIGMLGNASTGYGKIWLDDVSCDGEDSLMSC 662
Db 281 EWGTTICDDGWSYDAAVACKQLGCPATAICRVNASKGFGHILWDSVSCQGHFAVWQC 340
Qy 663 RNSGWNDCSHSEDDVGVICSDADMEIRLVGSGSRACAGKVEVNVQGVAGILCANGWGN 722
Db 341 KHEWGHKYNHEDAGVTCSDGSDLELRKRGSGRCAGTVEVEIQRLRLGKVCDDRGWGLK 400
Qy 723 LAEYVCRQLECGSALRVSRPHFTERTLHILMSNGCTGGASLWDCILRWKWKQTACHLN 782
Db 401 EADVCRQLGCGSALKTSYQVYSKIQATNTWLFSSCNGNETSLWDCNKGWGLTCDHY 460
Qy 783 MEALICSARHQPRLVGADMPGSRVYKVDHATWRSVCDSDFSLHAANVLCRELNCDAI 842
Db 461 BEAKITCSAHPRELVGGDIPCSGRVYKVDHATWRSVCDSDFSLEAASVLCRELQCGTW 520
Qy 843 SILVGDHFGKNGLTWAEKFCOEGSETHALCPVQHEDTCIHSREVGVCSRYTDVRL 902
Db 521 SILGAHFGEGNGQIWAEEFCOEGSETHALCPVAPRPEGTCSHSRDVGVCSTYETRL 580
Qy 903 VNGKSCDQGVQVEINVLGHWSGLCDTHWDPEDARVLCRLSCGTALSTGGKYIGERSVRV 962
Db 581 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKNGQI 640
Qy 963 WGRPHCLGNESLNDNCOMTVLGAPPCIHGNTVSVICTGSLTQPLPCLANVSDPYLSAV 1022
Db 641 WRHMFHCTGTQEHMGDCPVTALGASLCPSEQVAVSICSGNSQTLSSCNSSSLGTRPTI 700
Qy 1023 PEGSALICLEDKRLRLVGDGRSACGRVEIYHDGFWGTTICDDGWLSDAHVVCQKLGCGVA 1082
Db 701 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGWSGTTICDDSWLSDAHVVCRLGCGEA 760
Qy 1083 FNATVSAHFEGSGPILWDLNCTGESHLMQCPGSRGQHDCHKEHDAGVICSFTALR 1142
Db 761 INATGSAHFEGSGPILWDLNCTGESHLMQCPGSRGQHDCHKEHDAGVICSFTALR 820
Qy 1143 LYSETETESACRLEVFYNGTWGSGVRNITTAIAGIVCRLGCGENGVSAPLSKGTGS 1202
Db 821 LTSEASREACAGRLVFNAGWTVGKSSMSETTVGVVCRQLGCAKDKINPASILDKAMS 880
Qy 1203 GFWMVDDIQCPTKTHISIWQCLSAWERRISSPAETWTTCEDRIVRGDGTCSGRVFIW 1262
```

Db	341	KHEWGHKHYCNHNEDAGVTCDSGSDLELRLRGGGSRCACTVEVEIQRLLLGVKCDRGWGLK	400
Qy	723	IAEVVCRQLECGSAIRYSREPHEFTEFTHILMSNSGCTGGASLWDCIRWEMKOTACHLN	782
Db	401	EADVCRQLECGSALKTSQVYSKIQATYTWLFLSCNGNETSLWCKRWOGGLTCDHY	460
Qy	783	MEASLICSARHQRLPGADMPGSGRVEVYKHADTWRSVCDSDPSLHAANYCLRELNCGD	842
Db	461	EEAKITCSAHREPLRGDIPICSGRVEVYKHGDTWGSICDSDPSLEAASVLCRELQCGTVV	520
Qy	843	SLSVGDHFGKNGLTWAEKFOCGSETHALACPVIQHPEDTCHHSREVVYVCSR-----	896
Db	521	SILGAHFGGNGOIAEEFQCEGHESHLSCVPAPRPGTCSHSRDVGVYVSSKTOKTS	580
Qy	897	-----YTDRVLVNGKSCDQGOVEINVLGHWSLCLDTHW	929
Db	581	LIGSYTVKGTGLGSHSCLFLKPCLLPCYTEIRLVNGKTPCEGRVELKTLGANGSLCNSHW	640
Qy	930	DPEDARVLQRLSCGTALSTGTGKYIGERSVVRGHRFHC LGNESLLDNCQMTVLGAPPC	989
Db	641	DIEDAHLVCOOLKCGVALSTPGGARFGKNGQIWRHMFHCTGEQHMGDCEPVALGASLC	700
Qy	990	IHGNTSVICTGSLTQPLPCLANVSDPYLSAVPGSALICLEDKRLRLVDGDSRCAGRV	1049
Db	701	PSEQVASVICSNGQSTLSCNSLGPTRPTIPESAVACTESQRLRLVNGGRCAGRV	760
Qy	1050	EIYHDGFWGTICDDGWDLSDAHVVYCKLGGVAFNATVSAHFGEGSGPIWLDDLNCCTGE	1109
Db	761	EIYHGSWGTICDDSDWLSDAHVVYCRQLCGGAINATVSAHFGEGTGPWLDEMKCNKE	820
Qy	1110	SHLWQCPRGWGHDRHKEDAGVICSFTALRLVSETETESACARLEVFYNGTWGSVGR	1169
Db	821	SRIWOCHSHGWGQNCRHKEDAGVICSFTALRLVSETETESACARLEVFYNGTWGSVGR	880
Qy	1170	RNITTAGIVCRQLCGGNGVYSLAPLSTKSGPMWDDIDCPKTHLSIQCLSAPEH	1229
Db	881	SSMETTVGVVYCRQLCGGKIRNLSLADKAMSPMVDNVQCPGPTLQOCSSPWEK	940
Qy	1230	RISPAEETWITCEDIRVRGGDTESGRVETWHAGSWGTCDDSWDLAEAEVVCQQLGC	1289
Db	941	RLASPSEETWITCDNKLRLQEGTSCSGRVEIWHGSGNGTVCDDSWDLDDAQVVCQQLGC	1000
Qy	1290	GSALAAURLDASFGQGTGTIWLDDMRCKGNESFLWDCCHAKPWGQSDCGHKEDAGVCSQS	1349
Db	1001	GPALKAFKEAFEGGTGPIWLNEVVKCKGNESLWDCPARRWGRSCGHKEAANCTDIS	1060
Qy	1350	LK-----SLNASSGHALLTSIFFGLLLVLFII--LFTWCRVQKQKHLPLRVSTRRG	1401
Db	1061	VQKTPQKATGRSSROSSFIAGVILGVLLAIFVALFFLTKRRRRQ---LAVSSRGEN	1117
Qy	1402	SLEENLFHEMETCLKRED 1419	
Db	1118	LVHQIQYREMNLSCLNADD 1135	
RESULT	6		
ID	Q99MX8	PRELIMINARY; PRT; 1121 AA.	
AC	Q99MX8;		
DT	01-JUN-2001 (TReMBLrel. 17, Created)		
DT	01-JUN-2001 (TReMBLrel. 17, Last sequence update)		
DE	01-MAR-2002 (TReMBLrel. 20, Last annotation update)		
DE	Macrophage hemoglobin scavenger receptor CD163 precursor.		
GN	CD163		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
NCBI	TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21243729; PubMed=11345593;		
RA	Schaefer D.J., Boretta F.S., Hongegger A., Poehler D., Linnscheid P.,		
RA	Staege H., Muller C., Schoedon G., Schaffner A.,		

"Molecular cloning and characterization of the mouse CD163 homologue, a highly glucocorticoid-inducible member of the scavenger receptor cysteine-rich family."
RL Immunogenetics 53:170-177(2001).
DR EMBL: AF274883; AAK16065.1; -
DR MGD: MGI:2135946; Cdl63.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00530; SRCR; 9.
DR PRINTS; PR00258; SPERACTRCTR.
DR SMART; SM00202; SR; 9.
DR PROSITE; PS00420; SRCR_1; UNKNOWN_3.
DR PROSITE; PS0287; SRCR_2; 9.
KW Receptor; Signal.
FT SIGNAL 1 38 POTENTIAL.
SQ SEQUENCE 1121 AA; 120889 MW; 04428DF18AD26187 CRC64;

Query Match 40.2%; Score 3272; DB 11; Length 1121;
Best Local Similarity 52.4%; Pred. No. 6.9e-247;
Matches 582; Conservative 174; Mismatches 332; Indels 22; Gaps 9;
QY 320 GSDVYVLDGVSQGNESFLWDCRHSCTVNFDCLEHNDVSVICSDGADL--ELRLADGNN 377
Db 3 GHRMYLLAGAGSPGCKREVF---HLGFFVAVSLLSASAVTNAPGEMKKELRLAGGENN 58
QY 378 CSGRVEVRLEHQQWITCDONKNEQALVYVKGCPFFSVGSRRAKPSNEARDIWNIS 437
Db 59 CSGRVELKHDKWGTVCNSGMNNEVSVVCOOLGCPSTIKALGWANSAGSGYIMDKVYS 118
QY 438 CTGNESALWDCYTDGKAKRTCFRRSDAGVICSADLRLV-GAHSYCYGRLEVKYQGE 496
Db 119 CTGNESALWDCKHGDKGKHNCTHEKADAGVCTSDGNSLEMLVNSAGHCLGRVETKFOGK 178
QY 497 WCTVCHDRWSTRNAAVYVCKQCGKPMVFMFTYFKEASGPIWLDVSCIGNESINWDE 556
Db 179 WGTVCDDNFSDKHASVICKQCGSAISFGSAKLGAGSGPIWLDLACNGESALWDC 238
QY 557 HSGWKHNCVHREDVIVTCSGDATWGLRLVGSNRCRLEVFYQCRWGTVCDDGWSKA 616
Db 239 HSGWKHNCVHREDVIVTCSGDATWGLRLVGSNRCRLEVFYQCRWGTVCDDGWSKA 616
QY 617 AAVCSQDLPSSITGMGLNASTGYGKIWLDVSCDGDSDLSNCRSGMGNDCSHSE 676
Db 299 ASVVKQLGCPATSAIGRVNASESGQIWLNDNISCEGHEATLWCKHOENGKHYCHRE 358
QY 677 DVGVTCSADWELRLVGSRCRAGKVVNVGAGILCANGCMNTAEVVCRLQECGSA 736
Db 359 DAGVTCSDGADLELRLVGGSGRAGIVEVEIQKLTGKRCRSGWTLADADVVCRLQCGSA 418
QY 737 IRVSREPHTERTLHILMSNGCTGGEASLWDCIRWEKQTAACHLNMEASLICSAROPR 796
Db 419 LQTAQKIYKGTATWLFPGSCNGNETTFWQCKNWQGLSCDNFEAKVTCSGHREPR 478
QY 797 LVGADMPGSGRVEYKHADTWRSVCDSPSLHAANVLCRELNCGBAISLSVGDHFGKGNL 856
Db 479 LVGGEIPGSGRVEYKHADTWRSVCDSPSLHAANVLCRELNCGBAISLSVGDHFGKGNL 856
QY 857 TWAEKFOCGSETHALCPIVQHPEDTCHHSRVEGVVCSRYTDVRLVNGKSCDGOVEIN 916
Db 539 TWAEKFOCGSETHALCPIVQHPEDTCHHSRVEGVVCSRYTDVRLVNGKSCDGOVEIN 916
QY 917 VLGHWSLCDTHWDPEDARVLCRLSCGTALSTTGGKYIGERSVRVWGHFRHFCILGNESLL 976
Db 599 TLGAWGPLCSHWDHDAHLVLCQLKCGVAQSIPGEAFHFGKAGGVQVSHMFCHTGTTEHI 658
QY 977 DNCQMTVLGAPCTHGTNTVSVICTSLTQPLFPC--LANVSDPYLSAVPEGSALICLEDK 1034
Db 659 GDCLMTALGAPCTSEGQVAVSICSGNSQTLPCSSLSPVQTT--SSTIPKESEVPCTASG 717
QY 1035 RLRLVDGSRCAGRVEIYHDCFTWCTICDDGWDLSDAHVVCKQLGCGVAFNATVSAHREG 1094
Db 718 QLRLVGGRCAGRVEYCHESGWTVCDDNDWMDTANVVCRLQDCCGVAINATGSAFEG 777
QY 1095 SGPILWDDNCTGTESHLMQCPSPRGWGHDCRHKEDAGVICSEPTALRLXSETETESCAG 1154

Db 778 AGAIWLDEVICTGKESHVQCHSHGWRHCRKHKEDAGVICSEFMSRLTNEAHKENC 837
QY 1155 RLEVFYNGTWSVGRNRNITTAIGVIVCQLGGGNGVVSPLASTKSGFGFMVDDIOCPK 1214
Db 838 RLEVFYNGTWSVGRNRNITTAIGVIVCQLGGGNGVVSPLASTKSGFGFMVDDIOCPK 897
QY 1215 THISWQCLSPAPWERRISSP-AEETWITCEDRIRVRGDTGTCSGRVEIWHAGSWGTCVDD 1273
Db 898 GVDILWQCPSPFWKORQASPSQSWIICDNKIRIQEGHTDCSGRVEIWHAGSWGTCVDD 957
QY 1274 SWDLAEAEVVCQQLGCGSALALRDASFGQGTGTTLWDMRCCKGNESFLWDCCHAKPWQS 1333
Db 958 SWDLNDKAVVCKQLGCGQAVKALKEAFAFGPTGPIWLNKRCGNESFLWDCCHAKPWQS 1017
QY 1334 DCGHKEDAGVRCQSLKSLNASSGHLALILSSIFGLLLVILFLFTWC-----RVQKOK 1389
Db 1018 DCGHKEDASIOCLPKMTSESHGHTGHTPTALLVCGAILLVLLVILFLFTLKRRIQR-- 1075
QY 1390 HPLRVSTRRRGRSLEENLFHEMETCKRKED 1419
Db 1076 --LTVSSRGEVLHVQYQEMDS--KADD 1100

RESULT 7
QYUGM2 ID QYUGM2 PRELIMINARY; PRT: 2403 AA.
AC QYUGM2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE DMBT1/8kb.1 protein.
GN DMBT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RX MEDLINE=20065089; PubMed=10597221;
RA Mollenhauer J., Holmskov U., Wiemann S., Krebs I., Herberitz S.,
RA Madsen J., Kioschis P., Coy J.F., Poustka A.;
RT "The genomic structure of the DMBT1 gene: evidence for a region with
RT susceptibility to genomic instability";
RT Oncogene 18:6233-6240(1999).
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL: AJ243224; CAB63942.1;
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00530; SRCR; 14.
DR Pfam; PF00100; zona_pellucida; 1.
DR PRINTS; PR00258; SPERACTRCTR.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00202; SR; 14.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS00420; SRCR_1; UNKNOWN_13.
DR PROSITE; PS0287; SRCR_2; 14.
DR PROSITE; PS00682; ZP_DOMAIN; UNKNOWN_1.
FT CHAIN 26 2403 DMBT1/8KB.1 PROTEIN.
SQ SEQUENCE 2403 AA; 259573 MW; 9026222DE7399AD38 CRC64;

Query Match 36.7%; Score 2990; DB 4; Length 2403;
Best Local Similarity 40.5%; Pred. No. 2e-224;
Matches 619; Conservative 177; Mismatches 504; Indels 230; Gaps 23;

QY 44 TDLERLVNGDGPSCGVTEVVKFGQGWTCDDGWNNTASTVCKQLGCPFSFAM---FRF 100
Db 230 SSLALRLVNGDGRGRVREVLRYGWSGTCDDYWDYDNDANVVCRLQCGWAMSAPGNAQF 289

```
Qy 101 GOAVTRHGKWLDDVSCYNESALWECOHREWGSHNCHYGEDVGVNCYG----- 149
Db 102 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 150 -----EANGLRLVDGNNCSGRVEYKFEQERWGTICDDGNLNTAAVVCRLQC 198
Db 151 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 347 WPTSHASTAGPSSALRLVNGGDRGCRVEVLYRGSWGTVCDDSDWTSDANVVCRLQC 406
Db 348 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 199 PSSFISGVNPAVLRPIWDDTILCOGNELALNCHRRHGNHDCSHNEIDVTLCYDS- 257
Db 200 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 407 GWATSAFAGNARFGGSGPIVLDVRCSGYESYLWSPHNGWLTHNCHGSHEDAGVICS 466
Db 408 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 258 -----SDLELRVLGGTNCRMGVRLKIQGRWGTVCCHHKWNNAAADVCKQLG 304
Db 259 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 467 PTITLPASTVGSSESLALRLVNGGDRGCRVEVLYRGSWGTVCDDSDWTSDANVVCRLG 526
Db 468 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 305 CGTALHFAGLPHLOSGDVWLDGVCSGNPSFLWDCRHSTVNFDCILHONDVSVICS-D 363
Db 306 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 527 CGWAMLAPGNARFGGSGPIVLDVRCSGNPSFLWDCRHSTVNFDCILHONDVSVICS 586
Db 528 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 364 GADLELRVLGGTNCRMGVRLKIQGRWGTVCCHHKWNNAAADVCKQLG 423
Db 365 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 587 ESSLALGVNGGDRGCRVEVLYRGSWGTVCDDSDWTSDANVVCRLCGGATAPGNAR 646
Db 588 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 424 PSNEARDIWNISICTGNESALMDCTYDGAKARTCFRRSDAGVICS----- 469
Db 425 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 647 FGQSGPIVLDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 706
Db 648 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 470 -----DKADLDLRLVGAHSPCYGRLEVKYQGEWGTVCDBRWSTRNAAVVCRLQC 519
Db 471 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 707 TITLPPSTVGSSESLALRLVNGGDRGCRVEVLYRGSWGTVCDDSDWTSDANVVCRLQC 766
Db 708 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 520 GKPMHVFQMTYKKAASGPIWDDTILCOGNELALNCHRRHGNHDCSHNEIDVTLCYDS 576
Db 521 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 767 GWATSAFAGNARFGGSGPIVLDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICS 826
Db 768 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 577 -----GLRLVGSNRCGRLEVIYFQGRWGTVCDDGWNKAA 617
Db 578 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 827 SRPTSPDTWPTSHASTAGPSSALRLVNGGDRGCRVEVLYRGSWGTVCDDSDWTSDA 886
Db 828 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 618 AVVCSQLDCPSIITGMGLNASTGYKIWLDDVSCDGDDESILWSCRNSGWNNDCHSED 677
Db 619 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 887 NVVCRRLGCGWATSAFAGNARFGGSGPIVLDVRCSGYESYLWSPHNGWLTHNCHGSH 946
Db 888 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 678 VGVICSDA-----SDWELRLVGGSRCAKGVVNVQVAGI 713
Db 679 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 947 AGVICSAAHWSPTSPDTLPTITLPASTVGSSESLALRLVNGGDRGCRVEVLYRGS 1006
Db 948 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 714 LCANGWGNIAEVVVCRLQCGSAIRVSRPHFTERTLHILMSNGCTGGEASLWDCIRWE 773
Db 715 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 1007 VCDSDWTSDANVVCRLQCGWATSAFAGNARFGGSGPIVLDVRCSGHESYLWSPHNG 1066
Db 1008 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 774 WKQTACHLNMPEASLICSNAH-OP-----RUVGADMPCCSGRVE 809
Db 775 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 1067 WLHNCGHSESDAGVICSASQSRPTSPDTWPTSHASTAGPSSALRLVNGGDRGCRVE 1126
Db 1068 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 810 VKHADTRSVCDSDFLHAANVLRELNCDAISLSVGDHFKGNGLTWAELKFCQESGT 869
Db 811 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 1127 VLYRGSWGTVCDDSDWTSDANVVCRLQCGWATSAFAGNARFGGSGPIVLDVRCSGHES 1186
Db 1128 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 870 HLALCPVQHPEDTCIHSREGVVCSRYTD-----VRLVNG 905
Db 871 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 1187 YLWSPHNGWLTHNCHGSHEDAGVICSASQSRPTSPDTWPTSHASTAGPSSALRLVNG 1246
Db 1188 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 906 KSQCDGQVEINVLHGWSLCTHWDPEARVLCRLQCGTALSTGTGGYIGERSVRYWGH 965
Db 907 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 1247 GDRGCRVEVLYRGSWGTVCDDSDWTSDANVVCRLQCGWATSAFAGNARFGGSGPIVLD 1306
Db 1248 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 966 RFHCLGNESLLDNQMTVLGAPPCIHGNTVSVICTGSLTQPLFCLANVSDPYLSAVEP 1025
Db 967 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 1307 DVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSASQSRPT-PSPTWPTSHASTAG 1365
Db 1308 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
Qy 1026 SALICLEDKRLRLVDGSRCAGRVEIYHDFGWTICDDGWDLSDAHVVVCRLQCGVA 1085
Db 1027 GQVIVLDDVRCSGHESYLWSPHNGWLTHNCHGSHEDAGVICSAPLSRPTSPDT 346
```

```
Db 1366 SSL-----ALRLVNGGDRGCRVEVLYRGSWGTVCDDSDWTSDANVVCRLQCGWATSA 1419
Qy 1086 TVSAHFEGGSGPIWDDTILCOGNELALNCHRRHGNHDCSHNEIDVTLCYDS- 1135
Db 1420 FGSARFGGSGPIWDDTILCOGNELALNCHRRHGNHDCSHNEIDVTLCYDS- 1479
Qy 1136 -----SEFT-ALRLYSETETESACGRLEVFYNGTWGVRNRNITTAIAGIVC 1181
Db 1480 FDTWPTSRASSTAGSESTLALRLVNGGDRGCRVEVLYRGSWGTVCDDSDWTSDANVVC 1537
Qy 1182 ROLCGGNGVYSLAPLSK---TGSFPMVDDIOCPKTHISWQCLSAF- 1227
Db 1538 ROLGCG---GWAMSAFAGNAQFGGSGPIVLDVRCSGHESYLWSPHNGWLTHNCHGSHED 1594
Qy 1228 -----ERRISSPAEETWIT-----CEDRIRVR---GGDTCSRGEVIEWHAGSWGT 1269
Db 1595 GVICSAAQSQTSPRPTWLTNLPALTIVGSESSALRLVNGGDRGCRVEVLYRGSWGT 1653
Qy 1270 VCDSDWDLAEAEVVCQQLGCGSALALRDSAFGQGTGTIWLDDMRCKNESFLWDCCHAKP 1329
Db 1654 VCDSDWTSDANVVCRLQCGWATSAFAGNARFGGSGPIVLDVRCSGHESYLWSPHNG 1713
Qy 1330 WQSDCGHKGEDAGVRCSCQSLKLNASSGH 1359.
Db 1714 WLTHNCGHEDAGVICSATQINSTTTDWWH 1743

RESULT 8
Q9UGM3
ID Q9UGM3 PRELIMINARY; PRT; 2426 AA.
AC Q9UGM3
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE DBET1 prototype precursor.
GN DBET1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20065089; PubMed=10597221;
RA Mollenhauer J., Holmskov U., Wiemann S., Krebs I., Herberich S.,
RA Madsen J., Kioschis P., Coy J.F., Poustka A.;
RT "The genomic structure of the DBET1 gene: evidence for a region with
RT susceptibility to genomic instability."
RL Oncogene 18:6233-6240(1999).
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AJ243211; CAB63941.1;
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00530; SRCR; 14.
DR PRINTS; PR00100; zona_pellucida; 1.
DR SMART; SM00258; SPERACTRCTR.
DR SMART; SM00202; SR; 14.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS00420; SRCR_1; UNKNOWN_13.
DR PROSITE; PS00287; SRCR_2; 14.
DR PROSITE; PS00682; ZP_DOMAIN; UNKNOWN_1.
KW signal.
FT CHAIN 1 25 POTENTIAL.
FT CHAIN 26 2426 DBET1 PROTOTYPE.
SQ SEQUENCE 2426 AA; 262052 MW; 5A58FBC076FB7247 CRC64;

Query Match 36.7%; Score 2987.5; DB 4; Length 2426;
Best Local Similarity 39.7%; Pred. No. 3.3e-224;
Matches 615; Conservative 181; Mismatches 498; Indels 256; Gaps 22;
```

Qy	937	LCROLSCGTALSTGKGKYGISERSVRVWGHRRFHICIGNESLLDNCOMTVLGAPCPIHGNTVS	996
Dy	1172	VCROLGCGWAMSPAGNARFCQSGGPITLDDVRCSCGHESYLWSCPHNGWLSHNCGHEDAG	1231
Qy	997	VICTGSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVGDSCRAGRVEIYHDGF	1056
Dy	1322	VICASOSQPT-FSPDTWPTSHASTAGSESSL-----ALRLVNGDRCQGRVEVLYRGS	1284
Qy	1057	WGTCIDGDWDLSDAHVVCOKLGGVAFNATVSAHFGEGSGPIWLDDNLCTGTESHLMQCP	1116
Dy	1285	WGTVCDDYDWTNDANVVCROLGCGWATSAPGNARFGQSGPIVLDDVRCSCGHESYLWSCP	1344
Qy	1117	SRGNGHQDCHKEDAGVICS-----EFTALRLYSETETESC	1152
Dy	1345	HNGWLSHNCGHEDAGVICSASOSQPTSPDPWPTSHASTAGSESSLALRLVN--GGDRC	1402
Qy	1153	AGRLVEFYNGTWGSGVGRNRNITTAIGIYVCROLGCGENGWVSLAPLSK---TGSQFMWVDD	1209
Dy	1403	QGRVEVLYRSGWGTVCDDYDWTNDANVVCROLGCG---GWATSAPGNARFGQSGPIVLDD	1459
Qy	1210	IQCPKTHISTWQCLSPW-----ERRISSPAETWIT-----CEDR	1245
Dy	1460	VRCSCGHESYLWSCPHNGWLSHNCGHEDAGVICSASOSQPTSPDPWPTSHASTAGSEST	1519
Qy	1246	IRVR--GGDTESGRVEIWHAGSWGTCDDSDWLAEAEVVCQOLGCGSALAAALRDASFQ	1302
Dy	1520	LALRLVNGD-RCRGRVEVLYQSGWGTVCDDYDWTNDANVVCROLGCGWAMSPAGNAQFG	1578
Qy	1303	QGTGTIWLDDMRCKGNESFLWCHAKPQSGDCGHKEDAGVRCSGOSLKS	1352
Dy	1579	QGSQPIVLDDVRCSCGHESYLWSCPHNGWLSHNCGHEDAGVICSAAQSGS	1628

RESULT 9

Q9UJ57

PRELIMINARY;

PRT; 2412 AA.

AC	Q9UJ57	
DT	01-MAY-2000	(TEMBLrel. 13, Created)
DT	01-MAY-2000	(TEMBLrel. 13, Last sequence update)
DT	01-JUN-2002	(TEMBLrel. 21, Last annotation update)
DE	DMBT1/8kb.2	protein precursor.
GN	DMBT1.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
CX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=TRACHEA.	
RX	MEDLINE=99415938; PubMed=10485905;	
RA	Holmskov U., Mollenhauer J., Madsen J., Vitved L., Groenlund J.,	
RA	Tornøe I., Kilem A., Reid K.B., Poustka A., Skjodt K.;	
RT	"Cloning of gp-340, a putative opsonin receptor for lung surfactant	
RT	protein D.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 96:10794-10799(1999).	
CC	-!- SIMILARITY: CONTAINS 2 CUB DOMAINS.	
DR	EMBL; AJ243212; CAB56155.1; -.	
DR	HSSP; P29392; 1SFP.	
DR	InterPro; IPR000859; CUB_domain.	
DR	InterPro; IPR001507; Endoglin/CD105.	
DR	InterPro; IPR001190; Sscr_receptor.	
DR	Pfam; PF00431; CUB; 2.	
DR	Pfam; PF00530; SRCR; 14.	
DR	Pfam; PF00100; zona_pellucida; 1.	
DR	PRINTS; P00258; SPERACTRCPTR.	
DR	SMART; SM00042; CUB; 2.	
DR	SMART; SM00202; SR; 14.	
DR	SMART; SM00241; ZP; 1.	
DR	PROSITE; PS01180; CUB; 2.	
DR	PROSITE; PS00420; SRCR_1; UNKNOWN_13.	
DR	PROSITE; PS50287; SRCR_2; 14.	
DR	PROSITE; PS00682; ZP_DOMAIN; UNKNOWN_1.	
KW	Signal.	

```

Db 1187 DDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSTPPSPDTWPTSHASTAGSE 12448
QY 900 - - - - VRLVNCKSCDQGVETNVLGHWSLCDTHWDPEDARVLRCOLSCGTALSTGGKYI 955
Db 1247 SSLALRVLVGGDRCQGRVEVLYRGSNGTVCDDYTWOTNDANVVCRLQCGWATSAFGNARF 13008
QY 956 GERSVRVMGRHFHCLNGESLLDNQCMQTVLCAPPCIHGNTVSVICTGSLTQPLFPCLANVS 10151
Db 1307 QGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVICSASQSTP - PSPDTWP 13658
QY 1016 DPYLSAVPEGSALICLEDKRLRLVLGDGSRCAGRVEIYHDBGFTGCTDDGMDLSDAHVVCQ 107957
Db 1366 TSHASTAGSESSL - - - - -ALRLVGGDRCQGRVEVLYRGSNGTVCDDYWDNDANVVCRL 14151
QY 1076 KLGCGVAFNATVSAHFGBSGPTLWDDLNCTGTESHLMWCPGRGNGQHDCHRIKEDAGVIC 11338
Db 1420 QLGGWATSAFGNARFQGGSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEDAGVIC 14751
QY 1136 SEF - - - - -TALRLYSEFETESCAGRLVEFYNGTWSVGVRN 11751
Db 1480 SAFQSQSTPPSPDTWPTSRASTAGSESTLALRLVN - -GGDRCRGRVEVLYQGSNGTVCDDY 1537
QY 1172 ITTAIGIVCRLQCGENGVSUAPLSK - - -TGSGPMWVDDTOCPKTHISIWOLCSAPW- 1222
Db 1538 WDTNDANVVCRLQGC - - -GWASAPNAQFQGGSGPIVLDDVRCSGHEPYLWSCPHNGWL 1594
QY 1228 - - - - -ERRISSPAEETWIT - - - - -CEDRIRVR - - -GDGTSCSRV 1255
Db 1595 SHNCGHHEDAGVICSAAQSOSTPRDPTWLTNLPALTVCSSESLALRVNGCD -RCRGRV 1655
QY 1260 EIHAGSWGTVCCDDSWDLAEAYVVCQOLGCGSALAAALRDAFQGGTGTIWLDDMRCKNGE 1319
Db 1654 EVLYRGSWGTVCCDDSWDTNDANVVCRLQCGGWASAPGNARFQGGSGPIVLGDVRCSGNE 1711
QY 1320 SFLWDCHAKPWGSDGCHKEDAGVRCSGQSLKSLNASSGH 1359
Db 1714 SYLWSCPHKGLWLNHNCGHEDAGVICSATQINSTTTDWWH 1753

```

```

RESULT 10
ID Q96DU4 PRELIMINARY; PRT; 2413 AA.
ID Q96DU4;
AC 01-DEC-2001 (T-EMBLrel. 19, Created)
AT Q96DU4;
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE DBMT1/8kb.2 protein precursor.
DB DBMT1.
GN DBMT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE;
RA Mollenhauer J.;
RT "Major subforms of DBMT1 are gastrointestinal mucins that display
RT extensive alternative splicing and differential protein targeting."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RW EMBL: AJ297935; CAC44122.1; -.
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00530; SRCR; 14.
DR Pfam; PF00100; zona_pellucida; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS00420; SRCR.1; UNKNOWN_13.
DR PROSITE; PSS0287; SRCR.2; 14.
DR PROSITE; PS00682; ZP_DOMAIN; UNKNOWN_1.
DR PROSITE; PS00682; ZP_DOMAIN; UNKNOWN_1.
KW SIGNAL.
FT SIGNAL.

```


FT CHAIN 26 2413 DMPT1/8KB.2 PROTEIN.
SQ SEQUENCE 2413 AA; 260600 MW; 99A449C5F4F60728 CRC64;
Query Match 36.6%; Score 2982; DB 4; Length 2413;
Best Local Similarity 40.1%; Pred No. 8.7e-224;
Matches 618; Conservative 176; Mismatches 506; Indels 240; Gaps 23;
QY 44 TDLELRVNGDGCSTGVKVEFGQGWGTVDGWNNTASTVVKQLGCPFFSAM---PRF 100
DB 230 SSLALRVNGDRCGRVRLVYRGSGTVCDDYDWTNDANVVCRLGCGWAMAPGNAQF 289
QY 101 GOAVTRHGKIWLDDYSCYGNESALNECQREHNGSHNCYHGEVGYNCYV----- 149
DB 290 GOG---SGPIVLDVRCSHESYLWSCPHNGWLTHNCGHSESDAGVICSAPQSRRPTSPDT 346
QY 150 -----BANLGLRLVDGNNCSGRVEVKFQERWGTICDDGWNLTAAVVCRLGCG 198
DB 347 WPTSHASTAGPSSLALRVNGDRCQGRVEVLYRGSGTVCDDSDWTSDANVVCRLGCG 406
QY 199 PSFISGGVNVNPAVLRPIWLDDIILCOGNELALWCRHNGNHDPCSHNEDVTLFCY--- 255
DB 407 GWATSAPGNAREGQSGPIVLDVRCSGYESYLWSCPHNGWLSHNCQHSADAGVICSAAH 466
QY 256 -----DSSDLELRVNGTNRGMGRVELKTOGRWGTVCHHKWNA 294
DB 467 SRSTSPDTLPTITLPASTVGSSESLALRVNGDRCQGRVEVLYRGSGTVCDDSDWTN 526
QY 295 AADVCKQLGCGTALHFAFLHQLGSDVWVLDGYSCSGNESFLWDCRHSGTVPNDCILHQ 354
DB 527 DANVVCRLGCGWAMAPGNAREFGQSGPIVLDVRCSGNESYLWSCPHNGWLSHNCGHS 586
QY 355 NDVSYCS-DGADLELRADGNNCSGRVEVRIHQWTTICDQWKNQALVCKQLGCP 413
DB 587 EDAGVICSPESSALRVNGDRCQGRVEVLYRGSGTVCDDSDWTNDANVVCRLGCG 646
QY 414 FSVFGSRRAKPSNEARDIWINISCTGNESALWDCITYDGAKRKTFRSDAGVICS---- 469
DB 647 WATSAPGNAREGQSGPIVLDVRCSHESYLWSCPHNGWLSHNCGHEDAGVICSAAQS 706
QY 470 -----DKADLDLRVLAHSPCYKVEKYQGWGTVCHDRWSTRN 509
DB 707 RSTPGPDTLSTITLPPSTVGSSESLTLRVNGSDRCQGRVEVLYRGSGTVCDDSDWTND 766
QY 510 AAVVCCKQLGCGPMHVFMTGVEKASGPIWLDVYSCIGNESINWDCESHGKNCVHRE 569
DB 767 ANVVCRLGCGWATSAPGNAREFGQSGPIVLDVRCSHESYLWSCPHNGWLSHNCGHE 826
QY 570 DVIVTCS-----GDATW-----GLRLVGSNRCGRLEVYFQGRWGTV 607
DB 827 DAGVICSYSQSRPTSPDTWPTSHASTAGSESLALRVNGDRCQGRVEVLYRGSGTV 886
QY 608 CDDGNNRAAAVVCSQLDCPSIIGMGLGNASTGYGKIWLDDVDCDSDSLWSCNSGW 667
DB 887 CDDSDWTSDANVVCRLGCGWATSAPGNAREFGQSGPIVLDVRCSGYESYLWSCPHNGW 946
QY 668 GNNDCSHSESDVGVICSDA-----SDMELRVCGSSRRCAGKV 703
DB 947 LSHNCQHSADAGVICSAAHSTSPDTLPTITLPASTVGSSESLALRVNGDRCQGRV 1006
QY 704 EVNVQAGVILCANGWGNIAEVVVCRLQECGSAIRVSRPHTERTLHILMSNGCTGGE 763
DB 1007 EVLYGSGSGTVCDDSDWTNDANVVCRLGCGWAMAPGNAREFGQSGPIVLDVRCSGHE 1066
QY 764 ASLWDCIRWEKQATACHLNMAEALICSNAH-OP-----RLVG 799
DB 1067 SYLWSCPHNGWLSHNCGHSESDAGVICSASQSRPTSPDTWPTSHASTAGSESLALRVN 1126
QY 800 ADMPCSGRVEVKHADTWRSVCDPSFLHAANVLCRELNCDAISLVGCDHGKNGLTWA 859
DB 1127 GDRCQGRVEVLYRGSGTVCDDYDWTNDANVVCRLGCGWAMAPGNAREFGQSGPIV 1186
QY 860 EKFOCEGSETHALCPVQHPEDTCIHSREVGVVCSRYTD----- 899

DB 1187 DVCRCSHESYLWSCPHNGWLSHNCGHEDAGVICSASQSPPTSPDTWPTSHASTAGSE 1246
QY 900 ---VRLVNGSKQCDGOVEINVLGHWSICDTHWDPEDARVLCRLSCGTALSTTGGKYI 955
DB 1247 SSLALRVNGDRCQGRVEVLYRGSGTVCDDYDWTNDANVVCRLGCGWATSAPGNARF 1306
QY 956 GERSVYVWGHFRHCLGNESLLDNQCOMTVLGAAPPCTIHTGNTVSVICTGSLTQPLFPCLANVS 1015
DB 1307 GQSGPIVLDVRCSGHESYLWSCPHNGWFSHNCGHEDAGVICSASQSPPT-PSPDTPW 1365
QY 1016 DPLYSAPVEGSALECLDKRLRLVDGDSRCRAGRVEIYHDGFWGTICDDGWLSDAHVVCQ 1075
DB 1366 TSHASTAGSESL-----ALRLVNGDRCQGRVEVLYRGSGTVCDDYDWTNDANVVC 1419
QY 1076 KLGCQVAFNATVSAHFEGSGPIVLDLNCCTGTHLWQCPSRGWOHDCRHKEDAGVIC 1135
DB 1420 QLGCQWATSAPGNAREFGQSGPIVLDVRCSGHESYLWSCPHNGWLSHNCGHEDAGVIC 1479
QY 1136 -----SEFT-ALRLYSETETESCAGRLVEFYNGTWGSGVRGN 1171
DB 1480 SASQSPPTSPDTWPTSPRSTAGSESLALRVN--GGDRCRGRVEVLYQSGWTVCCDY 1537
QY 1172 ITTAITAGIVCRLGCGGNGVVS LAPSK---TSGSFMWDDIQCPKTHISTWQCLSAW- 1227
DB 1538 WDTNDANVVCRLGCG---GWAMAPGNAQFGQSGPIVLDVRCSGHESYLWSCPHNGWL 1594
QY 1228 -----ERRISSPAEETWIT-----CEDRIRVR---GGDTECSGRV 1259
DB 1595 SHNCGHEDAGVICSASQSPPTPRDPTWLTNLPALTVGSSESLALRVNGD-RCRGRV 1653
QY 1260 EIWAGSGTVCDDSDWLAEAEVVCQGLGCGSALALRADASFGOGTGTWILDDMRCKGNE 1319
DB 1654 EVLYRGSGTVCDDSDWTNDANVVCRLGCGWAMAPGNAREFGQSGPIVLDVRCSGNE 1713
QY 1320 SFLNDCHAKPMGQDCGHKEDAGVRCQSLKSLNASSGH 1359
DB 1714 SYLWSCPHKGLTHNCGHEDAGVICSATQINSTTTDWWH 1753

RESULT 11

Q9UKJ4 PRELIMINARY; PRT; 2413 AA.
ID Q9UKJ4 AC Q9UKJ4
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Gp-340 variant protein.
GN DMBT1
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=9415938; PubMed=10485905;
RA Holmskov U., Mollenhauer J., Madsen J., Vitved L., Gronlund J.,
RA Tjorne I., Kliehm A., Reid K.B., Poustka A., Skjold K.;
RT "Cloning of gp-340, a putative opsonin receptor for lung surfactant
protein D.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:10794-10799(1999).
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AF159456; AAD49696.1; -;
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00431; CUB; 2
DR Pfam; PF00530; SRCR; 14
DR Pfam; PF00100; zona_pellucida; 1.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00202; SR; 14.
DR SMART; SM00241; 2P; 1.
DR PROSITE; PS01180; CUB; 2.

DR EMBL; AB020821; BAA78577.1; JOINED.
DR EMBL; AB020822; BAA78577.1; JOINED.
DR EMBL; AB020823; BAA78577.1; JOINED.
DR EMBL; AB020824; BAA78577.1; JOINED.
DR EMBL; AB020825; BAA78577.1; JOINED.
DR EMBL; AB020826; BAA78577.1; JOINED.
DR EMBL; AB020827; BAA78577.1; JOINED.
DR EMBL; AB020828; BAA78577.1; JOINED.
DR EMBL; AB020829; BAA78577.1; JOINED.
DR EMBL; AB020830; BAA78577.1; JOINED.
DR EMBL; AB020831; BAA78577.1; JOINED.
DR EMBL; AB020832; BAA78577.1; JOINED.
DR EMBL; AB020833; BAA78577.1; JOINED.
DR EMBL; AB020834; BAA78577.1; JOINED.
DR EMBL; AB020835; BAA78577.1; JOINED.
DR EMBL; AB020836; BAA78577.1; JOINED.
DR EMBL; AB020837; BAA78577.1; JOINED.
DR EMBL; AB020838; BAA78577.1; JOINED.
DR EMBL; AB020839; BAA78577.1; JOINED.
DR EMBL; AB020840; BAA78577.1; JOINED.
DR EMBL; AB020841; BAA78577.1; JOINED.
DR EMBL; AB020842; BAA78577.1; JOINED.
DR EMBL; AB020843; BAA78577.1; JOINED.
DR EMBL; AB020844; BAA78577.1; JOINED.
DR EMBL; AB020845; BAA78577.1; JOINED.
DR EMBL; AB020846; BAA78577.1; JOINED.
DR EMBL; AB020847; BAA78577.1; JOINED.
DR EMBL; AB020848; BAA78577.1; JOINED.
DR EMBL; AB020849; BAA78577.1; JOINED.
DR EMBL; AB020850; BAA78577.1; JOINED.
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00530; SRCR; 9.
DR Pfam; PF00100; zona_pellucida; 1.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS00420; SRCR_1; UNKNOWN_8.
DR PROSITE; PS0287; SRCR_2; 9.
DR PROSITE; PS00682; ZP_DOMAIN; UNKNOWN_1.
SQ SEQUENCE 1785 AA; 193991 MW; 38B2363F95226EB0 CRC64;

Query Match
Best Local Similarity 30.1%; Score 2448; DB 4; Length 1785;
Matches 534; Conservative 156; Mismatches 413; Indels 374; Gaps 28;

QY 43 GTD--LELRVNGDPCSGTVEVFGQGWGTVCDDGNTTASTVVCQKLGCPFSFAM--- 97
DB 95 GSDGLALRLVNGDRCGRVEILYRGSGTVCDDSDWTDNDANVVCRLGCGWAMSAPGN 154

QY 98 FRFGQAVRHGKIWLDDVSCYGNESALWECOHREWSHCNCHYHGEDVGNYC----- 148
DB 155 AWFQGG---SFIALLDVRCSGHESYLWSCPHNGWLSHNCGHGEDAGVICSAAPOQSTLR 211

QY 149 -----GRANLGLRLVGNNSCGRVEVKFQERWGTICDDGWNINLTAAYV 192
DB 212 PESWPVRISPPVPTTEGSESSALRLVNGDRCGRVEILYRGSGTVCDDYWDNDANV 271

QY 193 CROLGCPSSFTSSGVNNSPAVLRIWLDDILCQNELALWNCRRHGNHDCSHNEDVTL 252
DB 272 CRLGCGWAMSAPGNAOFGQSGPIVLDDVRCSGHESYLWSCPHNGWLTNCHGSEDAGV 331

QY 253 TCY-----DSSDLRLVLVGTNRCMGRVCLKIOGRWGTVCCHK 290
DB 332 ICAPQSRPTSPDPTWPTSHASTAGSESSALRLVNGDRCGRVEILYRGSGTVCDD 391

QY 291 WNNAAADVVCQKLGCGTALHFAPLHQLSGSDVYVWLGVSCSGNESFLWPCRHSGTYNFD 350
DB 392 WDTSDANVVCRLGCGWATSAPGNARFGQSGPIVLDDVRCSGHESYLWSCPHNGWLSHN 451

QY 351 CLHONDVSVICSDGADLELRADLADGNNCSGRVEVRIHQWWTICQDNKNEQALVVCQKL 410
DB 452 CQHSEDAGVCSAA-----HSWT----- 470

QY 411 GCFFSVFGSRRAPSNARDIWINISICTGNESALWDCTYDGKAKRKTCFRRSDAGVICSD 470
DB 471 -----PSPDTLPITITLPASTVGSESS----- 491

QY 471 KADLDRLVLGAHSPCYGRLEVYOGEGWTVCHDRKSTNAAAVVCQKLGCGKPMHVFQMTY 530
DB 492 ---LALRLVNGDRCGRVEILYRGSGTVCDDSDWTDNDANVVCRLGCGWAMSAPGNAR 548

QY 531 FKEASGPIWLDDVSCIGNESNIWDCHEGSGKHCNCHVREDVLTVCSDGA-----TW- 581
DB 549 FGQSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHSEDAGVICSASQSRPTSPDPTW 608

QY 582 -----GLRLVGSNRCSGRLEVYFGRWGTVCDDGWNNSKAAAAYVCSQDLCPS 628
DB 609 TSHASTAGSESSALRLVNGDRCGRVEILYRGSGTVCDDYWDTDNDANVVCRLGCGW 668

QY 629 SLICMGLGNASTGYGIWLDDVSCDGEDSLWSCNNSGWNNDCHSHSDEDVGICS----- 683
DB 669 AMSAPGNARFGQSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHSEDAGVICSASQ 728

QY 684 -----DASDMELRLVGGSSRCAGKVEVNVQGVAVGILCANGWGNIAEV 726
DB 729 PTPSPDPTWPTSHASTAGSESSALRLVNGDRCGRVEILYRGSGTVCDDYWDTDNDANV 788

QY 727 VCRLGCGSATRSREPHFTERTLHILMSNGCTGGASLWDCIRWEMKQATACHLMEAS 786
DB 789 VCRLGCGWATSAPGNARFGQSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHEDAG 848

QY 787 LICSAHR-QPRLVGDMPGCSGRVEVKHADTWRSVCDSDSFLHAANVLCRELNCGDAISLS 845
DB 849 VICASASQ-----TPSPDPTWPT----- 867

QY 846 VGDHFGKNGLTWAEKFCGESETHALCPVQHPEDTCHSREVGVVCSRYTQVRLVNG 905
DB 868 -----SRASTAGSESTAL-----RLVNG 886

QY 906 KSCQDQGVENVLGHWSGLCDTHWDPEARVLCRLQSCGTALSTTGGKYIGERSVRVWGH 965
DB 887 GDCRCGRVEILYRGSGTVCDDYWDTDNDANVVCRLGCGWAMSAPGNAOFGQSGPIVL 946

QY 966 RFHCLGNESLIDNCOMTVLGAPCIHGNTVSVICTGLTO--PLFPCLANVSDPYLSAVP 1023
DB 947 DVRCSGHESYLWSCPHNGWLSHNCGHEDAGVICSASQSTPRPDTWLTNLPALTVG 1006

QY 1024 EGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVQCKLGCGVAF 1083
DB 1007 ESSL-----ALRLVNGDRCGRVEILYRGSGTVCDDSDWTDNDANVVCRLGCGWAM 1059

QY 1084 NATVSAHFQSGSGPIWLDDILCTESHLMQCPSRQGHQDCHKEADAGVICSFTALRL 1143
DB 1060 SAPGNARFGQSGPIVLDDVRCSGHESYLWSCPHNGWLTNCHGHDAGVICS---ATQI 1116

QY 1144 YSET-----ETESACGRLEVFY-NGTWGS----- 1166
DB 1117 NSTTDDWHPTTTTARPNSSCGFL--FYASGTFSSPYPAAYPNNAKCVWEIEVNSGY 1174

QY 1167 ---VGRNI-----TTAAGIVC---ROLGCGENVVSL---AP 1196
DB 1175 RINLGFNLKLEAHNCSFDYVEIFDGLNSLLILKICNDTRQIFTSSYNRMTIHRSD 1234

QY 1197 LSKTGSFM-WVDDIQCPKTHISIWQSLAPWERRISSPAEETWITCEDRIRVGGGTDEC 1255
DB 1235 ISQNTGFLAWYN-----SFPSDATL-----RLVNLNSSYGLC 1267

QY 1256 SGRVEIHWAGSWGTVCCDSDWDLAEAEVVCQKLGCGSALALRDASFCQGTCTIWLDDMRC 1315
DB 1268 AGRVEIYHGGTGTGTVCCDSDWTIOEAEVVCRLGCGRAVSALGNAYFGSGSGPITLDDVEC 1327

```
Qy 1316 KGNESFLWCHAKPMWGSDCGHKEDAGVRCSGQSLKS 1352
| | | | | : | : | | | | | | | | | | | | | |
Db 1328 SGTESTLWQCRNRGWFSHNCHNHRDAGVICSGNHLS 1364

RESULT 13
O97375 PRELIMINARY; PRT; 2153 AA.
AC O97375;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Scavenger receptor cysteine-rich protein type 12 precursor.
OS SCRI2.
GN Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Echinozoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99328904; PubMed=10398804;
RA Pancer Z., Rast J.P., Davidson E.H.;
RT "Origins of immunity: transcription factors and homologues of effector
RT genes of the vertebrate immune system expressed in sea urchin
RT coelomocytes.";
RL Immunogenetics 49:773-786(1999).
DR EMBL: AF064259; AAD08654.1; -.
DR InterPro: IPR001190; SRCR_receptor.
DR Pfam: PF00530; SRCR; 20.
DR PRINTS: PR00258; SPERACTRCPT.
DR SMART: SM00202; SR; 20.
DR PROSITE: PS00420; SRCR_1; 15.
DR PROSITE: PS00287; SRCR_2; 20.
KW Receptor; Signal.
FT SIGNAL 1 16
FT CHAIN 17 2153
FT TYPE 12.
SQ SEQUENCE 2153 AA; 226557 MW; 98C51996A272A010 CRC64;

Query Match 30.08; Score 2439.5; DB 5; Length 2153;
Best Local Similarity 38.28; Pred. No. 1.9e-181;
Matches 52; Conservative 197; Mismatches 560; Indels 97; Gaps 22;

Qy 32 LNSCFLISSFN-----GTDLELVLVNGDPCSGTVEVKFQGGTVCDDGWN 78
| | | | | : | | | | | | | | | | | | | | | |
Db 113 LSSCAVSSHNAHSEDAGVVCAGSDGNIRLVGSHSSEGRVELYVNNQNGTVCDDLWD 172
| | | | | : | | | | | | | | | | | | | | | |
Qy 79 TTASTVCKQLGC-PFSFAMPRFQAVTRHGKIWLDDVSCYGNESALWECOHREWGSHNC 137
| | | | | : | | | | | | | | | | | | | | | |
Db 173 LNDQVACRQLGLGPAVSADAHFEG---SGSILLDNVACITGTETSLSCSHPGTSHNC 229
| | | | | : | | | | | | | | | | | | | | | |
Qy 138 YHGEDVGVCYGEANGLRLVDGNNSCGRVEVKFQERWGTICDDGNLNTAAVVCROLG 197
| | | | | : | | | | | | | | | | | | | | | |
Db 230 GHSDAGVICSRRGSDIRVVGGSIPTEGRVEFVNGAWGTVCDDLWDINDASVACROLG 289
| | | | | : | | | | | | | | | | | | | | | |
Qy 198 CPSSFISGVVNSPAVLRIWLDDILCOGNELALWNCRHGRWGHDCSHNEDVTLTCYDS 257
| | | | | : | | | | | | | | | | | | | | | |
Db 290 FGRAISAGGASYQGSGSIFLDNLACTGAESNLLSCPHNGVSHNCHGEDAGVLCSSS 349
| | | | | : | | | | | | | | | | | | | | | |
Qy 258 SDLELRLVGGTNRMGVRELKIOGRWGTVCCHKNWNAADVVCVKQLCGGTALHFAGLPHL 317
| | | | | : | | | | | | | | | | | | | | | |
Db 350 LSSDIRVVGGRSPTEGRVEFVNRWGTVCDDLWDVNDTNVACRQLGFGRAISAPGGASY 409
| | | | | : | | | | | | | | | | | | | | | |
Qy 318 QSGSDVWLDCVSCSGNESFLWDCRHSGTGVNFDCLHQNDVSVICSDGA-DLELRLADGSN 376
| | | | | : | | | | | | | | | | | | | | | |
Db 410 QGSGSILLDNLACTGAESSLLSCPHNGVSHNCHGEDAGVSCAPSSQESRVRVLVGLN 469
| | | | | : | | | | | | | | | | | | | | | |
Qy 377 NCSGRVEIRIHEQWNTICDQWKNQEQALVCKQLGCPSPFVSGSRRAKP--SNEARDIWIN 434
| | | | | : | | | | | | | | | | | | | | | |
Db 470 NREGRVEIFLNQNGTVCDDHGTDPDANVVCRLQGLFSG--GSARSSAYFGSGVPILLD 527
| | | | | : | | | | | | | | | | | | | | | |
Qy 435 SISCTGNESALWDCTYDGAKARTCFRRSDAGVIC---SDKADLDLRLVGAHSPCYGRLEV 491
| | | | | : | | | | | | | | | | | | | | | |
```

DT 01-NOV-1999 (TrEMBIrel. 12, Last sequence update)
DE 01-JUN-2002 (TrEMBIrel. 21, Last annotation update)
DE DMBT1/6kb.1 protein precursor.
GN DMBT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RX MEDLINE=97434209; PubMed=9288095;
RA Mollenhauer J., Wiemann S., Scheurle W., Korn B., Hayashi Y.,
RA Wilgenbus K.K., von Deimling A., Poustka A.;
RT "DMBT1, a new member of the SRCR superfamily on chromosome 10q25.3-
RT q26.1 is deleted in malignant brain tumours";
RL Nat. Genet. 17:32-39(1997).
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL, AJ000342; CAA04019.1; -;
DR InterPro; IPR000859; CUB_Domain.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00530; SRCR; 9.
DR Pfam; PF00100; zona_pellucida; 1.
DR PRINTS; PR00258; SPRACTRCPTR.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00241; SR; 9.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS00420; SRCR_1; UNKNOWN 8.
DR PROSITE; PS0287; SRCR_2; 9.
DR PROSITE; PS00682; ZP_DOMAIN; UNKNOWN_1.
DR SIGNAL.
KW SIGNAL.
FT CHAIN 1 25 POTENTIAL.
FT CHAIN 26 1785 DMBT1/6KB.1 PROTEIN.
SQ SEQUENCE 1785 AA; 193941 MW; 7B1F8D47E4A82092 CRC64;

Query Match 29.94; Score 2435; DB 4; Length 1785;
Best Local Similarity 36.19; Pred. No. 3.3e-181;
Matches 533; Conservative 155; Mismatches 415; Indels 374; Gaps 28;

QY 43 GTPD--LELRVNGDPCSGTVEVFQGGWGTVCDDGWNTTASTVVCVKOLGCPESFAM--- 97
DB 95 GSDSLRLVNGDRCGRVEILFRSGWGTVCDDSWDTNDANVVCRLGCGWMSAPGN 154
QY 98 FRFGQAVTRHGKIWLDYSCYGNESALWECQHRWEGSHNCYHGEDVGYNCY----- 148
DB 155 AWFQGG---SGPIALDDVRCSGHESYLWSCPHNGWLSHNCGHGEDAGVCSAAQPSLRL 211
QY 149 -----GEANLGLRLVDGNNSCGRVEYKFOERWGTICDDGNLNTAAVV 192
DB 212 PESWVRISPPVPTGESLALRLVNGDRCGRVEYLRGSGWGTVCDDYWDNDANV 271
QY 193 CRLGCPSSFTSSGVNPAVLRLPWLDDILCOGNELALWNCRHRGWNHDCSHNEDVTL 252
DB 272 CRLGCGWMSAPGNAQFGQGGPIVLDVRCSGHESYLWSCPHNGWLSHNCGHEDAGV 331
QY 253 TCY-----DSSDLRLVGGTNRGMRVVELKIQGRWGTVCCHK 290
DB 332 ICSPQSRPTSPDPTWPTSHASTAGPESLALRLVNGDRCGRVEYLRGSGWGTVCDD 391
QY 291 WNNAAADVVCOLGCGTALHFAGLPHLOGSDVWLDGVCSGNESHFLWDCRHSCTVNF 350
DB 392 WDTSDANVVCRLGCGWATSAPGNARFGQGGPIVLDVRCSGHESYLWSCPHNGWLSH 451
QY 351 CLHONDVSCDGADLELRADGNSNCGRVEIRHQWTTICDQWKNQALVVCVKQL 410
DB 452 COHSEDAGVCSAA-----HSWST----- 470
QY 411 GCPFSVGRRAKPSNEARDIWNISICTGNESALWDCYDCKAKRTCTFRSDAGVCS 470
DB 471 -----PSPDTLPTITLPAVTSVSS----- 491

QY 471 KADLDLRLVGAHSPCYGRLEVYQGEWGTVCCHDRWSTNRNAAAVCKQLGCGKPMHVFMTY 530
DB 492 ---LALRLVNGDRCGRVEYLRGSGWGTVCDDSWDTNDANVVCRLGCGWMSAPGNAR 548
QY 531 FKEASGPVLDVSCIGNESINWCEHSGKHNCHVEDVIVTCSGDA-----TW- 581
DB 549 FGQGGPIVLDVRCSGHESYLWSCPHNGWLSHNCGHEDAGVICSASQSRTSPDTPW 608
QY 582 -----GLRLVGGNRCGRLEVYFOGRWGTVCDDGWNSKAAAVVCSOLDCPS 628
DB 609 TSHASTAGSESLALRLVNGDRCGRVEYLRGSGWGTVCDDYWDNDANVVCRLGCGW 668
QY 629 SIIGWGLNASTGYGKIWLDVSCDGBESDLWSCRNSGNWGNDCSHSDEYGVICS----- 683
DB 669 AMSAPGNARFGQGGPIVLDVRCSGHESYLWSCPHNGWLSHNCGHEDAGVICSASQS 728
QY 684 -----DASDMELRLVGGSSRCACKVENVVQGAAGVILCANGWGNIAEV 726
DB 729 PTPSPDPTWPTSHASTAGSESLALRLVNGDRCGRVEYLRGSGWGTVCDDYWDNDANV 788
QY 727 VCRQLEGCSALRVSRPHFTERTLHILMSNGCTGGEASLWDCIRWENKQOTACHLNMEAS 786
DB 789 VCRQLGCGWATSAPGNARFGQGGPIVLDVRCSGHESYLWSCPHNGWLSHNCGHEDAG 848
QY 787 LICSAHR-QPRLVGADMPGSRVEYKHADTWRSVCDSDFSLHAANVLCRELNCDAISLS 845
DB 849 VICASASQSP-----TPSPDPTWPT----- 867
QY 846 VGDHFGKNGLTWAEKFOEGSETHLALCPVQHPEDTCIHSREVGVCSRYTVDRLVNG 905
DB 868 -----SRASTAGSESLAL-----RLVNG 886
QY 906 KSQCDGQVEINVLGHWSGLCDTHMDPEDARVLCRLSCGTALSTGGKYIGERSYRVWGH 965
DB 887 GDRGRGRVEYLRGSGWGTVCDDYWDNDANVVCRLGCGWMSAPGNAQFGQGGPIVLD 946
QY 966 RFHCLGNESLLDNQMTVLGAPPCIHGNTVSVICTGSLTQ--PLFPCLANVSDPILSAPV 1023
DB 947 DVRCSGHESYLWSCPHNGWLSHNCGHEDAGVICSAAQSQTTPRDTWLTTLNLPALT 1006
QY 1024 EGSALICLEDRLRLVDGSDRCAGREYIHDGFWGTICDDGWLDSDAHVVCQKLGCGVAF 1083
DB 1007 ESSL-----ALRLVNGDRCGRVEYLRGSGWGTVCDDSWDTNDANVVCRLGCGW 1059
QY 1084 NATVSAHFGESGPVLDLNCGTGTEHLMOCPSRGWGHDCRHKEDAGVICSSETALRL 1143
DB 1060 SAPGNARFGQGGPIVLDVRCSGHESYLWSCPHNGWLSHNCGHEDAGVICS-----ATQI 1116
QY 1144 YSET-----ETESCAGRELEVY-NGTWGS----- 1166
DB 1117 NSTTDDWHPHTTTTARPSNCGGFL--FYASGTFSPSPYPAYPNNKAKVWEIEVNSGY 1174
QY 1167 ---VGRRI-----TTAGIVC---ROLGCGENGVSLL---AP 1196
DB 1175 RINGFSLNLEAHNCSFDYVEIFDGSINSLLLGKICNDTRQIFTSYNMTTHFRSD 1234
QY 1197 LSKTGSFPM-WVDIIQCPKTHISIWQCLASAPWERRISSPAETWTICEDRIVRGDTEC 1255
DB 1235 ISFQNTGFLAWN-----SPPSDATL-----RLVNLNLSYGLC 1267
QY 1256 SGRVEIWAGSWGTVCCDSDWDLAEAVVCOQLGCSALALRDASFGOGTGTIWLDDMRC 1315
DB 1268 AGREYIHHGGWGTVCDDSWTIQAEAVVVCRLGCGRAVSALGNAYFGSGSPITLDDVE 1327
QY 1316 KGNESFLWDCWQSDCGHKEDAGVRCSGQSLSK 1352
DB 1328 SGTESTLWQCRNRGWSHNCNHRDAGVICSGNHLST 1364

RESULT 15
O96943
ID O96943 PRELIMINARY; PRT; 2043 AA.

AC 096943;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE SRCR domain, membrane form 2.
GN SRCR2.
OS Geodia cydonium (Sponge).
OC Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
OC Astrophorida; Geodidae; Geodia.
OC NCBI_TaxID=6047;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98369060; PubMed=9701562;
RA Blumberg B., Pancer Z., Diehl-Seifert B., Steffen R., Muenkner J.,
RA Mueller I., Mueller W.E.G.;
RT "The putative sponge aggregation receptor: Isolation and
RT characterisation of a molecule composed of scavenger receptor
RT cysteine-rich domains and short consensus repeats.";
RL J. Cell Sci. 111:2635-2644(1998).
DR EMBL; Y14953; CAA75175.1;
DR HSP; P10998; 1VV0.
DR InterPro: IPR001190; Srcr_receptor.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00530; SRCR_14.
DR Pfam: PF00084; sushi; 6.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00032; CCP; 6.
DR SMART; SM00202; SR; 14.
DR PROSITE; PS00420; SRCR_1; 8.
DR PROSITE; PS0287; SRCR_2; 14.
SQ SEQUENCE 2043 AA; 220896 MW; 493319D0C62FAAE CRC64;

Query Match 27.6%; Score 2246.5; DB 5; Length 2043;
Best Local Similarity 32.4%; Pred. No. 2.1e-166;
Matches 514; Conservative 196; Mismatches 588; Indels 287; Gaps 31;

Qy 7 SWHIDFGCCCHQLFSAVVCILLNCLFLISSFNFGTDLRLV-NGDGPCSGTVEVKF 65
Dy 114 AWH-SVGSECSHQNNSGVCT-----DSSAVRLVEGDMNRRGRVEYH 156
Qy 66 QGQNGTVCDDGNTASTVVCVKLGCPSPFAMFRGQAVTRHGKIWLDDVSCYGNESALW 125
Dy 157 ADEWGSICDDYVFGGEAVVICHGLGFSAEDESHGGTTPQGDGLIWMNDLQCTHESIA 216
Qy 126 EQCHREMGSHCYHGEDGVNVCYGEAN-----LGLRLVDGNNCS-----GRVEYKQF 173
Dy 217 ECTPFGWGEHCNQREDAGVCTATHPDFTYIIRLNGTSSNATHGLQYGRVEVEIN 276
Qy 174 ERWGTICDDGNLNTAAVVCRLGCPSPFSISSGVVNSPAVLRL---PIWLDDILCOGNELA 230
Dy 277 NTWGTVCDDGNGIEDATACRLQGLFYGALLASSRSEYAOVAPSSVPIVWDEVLCGDEVA 336
Qy 231 LWNCRHRCWGHNDGSHNEDVTLTCYDSSDLE-----LRLVCGTNRCM-----GRVEL 277
Dy 337 LSECHHRTTGVHDCSHYEDAGVFC--TTNDIQPAAPKVLRLVNNNGSVDEGDGHVEGRVEV 395
Qy 278 KIOGRWGVCHHKWNAADVVCVKLGCGTALHFAGLPHLOSGDGVVWLDGVSCSGNESEF 337
Dy 396 NFYGVWGTICDDYWSLADGVICKMLGFTYADEGVYSGAFGIGTGPIMWDDVQCDGTET 455
Qy 338 LWDCHRGSTVNFCDLHONDVSVICS---DGADLELRLAGDNNCSGRVEYRIHEQWMTIC 394
Dy 456 IDHCDFFPGWGVHNCYHFPEDASVCSYRQPGSD---VRLVGGSSMBGRVEVLHNGVWGV 513
Qy 395 DONKNEQALVVCVKLGCPSPFSVFSRRAPK-----SNEARDIWNISCTGNESALWDC 448
Dy 514 DDSWGAEDAQVICHQLG----FYGTATAVRGTFQGFASSQPIWLDDVGCCTGSEMYLSDC 569
Qy 449 TYDGKAKRTCFRRSDAGVIC-----
Dy 570 SSSGFGNHCNYHYEDAGVGVCGTSPFAIPVRLSNGTRPNEGTVETIPYQNRWGTICDIFWTL 629
Qy 469 ----- 468

Search completed: May 12, 2003, 06:42:56

Job time : 71 secs

Pred. No. is the number of results predicted by chance to have a

TITLE

ALIGNMENTS

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

transmembrane molecule (M160) expressed by human macrophages
J. Immunol. 165 (11), 6406-6415 (2000)
20540099
11086079
2 (bases 1 to 4579)
Groenlund, J., Vitved, L., Lausen, M., Skjold, K. and Holmskov, U.
Direct Submission
Submitted (05-MAY-2000) Department of Immunology & Microbiology,
Institute of Medical Biology, University of Southern Denmark,
Odense University, Winsloe Parken 21.1, Odense 5000, Denmark
Location/Qualifiers
1. 4579
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/map="12q13.3"
1. .25
26. .4387
/codon_start=1
/product="scavenger receptor cysteine-rich type 1 protein
M160 precursor"
/protein_id="AAF91396.1"
/db_xref="GI:9652087"
/translation="MWLPQNSWHIDFGRCCHONLFSAVVTCILLNSCFRISFNCT
DLELRVNGDPCSCGVTKVKGQNGTVCDGWNVTASTVYCKQLCGCPFAFRFGQ
AVTRGKIWDVSCYGNESALWECQHRHSHNCYHGEDVGVNCGYGANLGLRLVDG
NNSCSRVEVPQERWGTICDDGNLNTAAVVCRLQGPSPFISGVVNSPAVLRIPI
LDLILCOGNEALNHRHGNHDSNEDVTLFYDSDDLRLVNGTNRGMRVE
LKIQGRWTVCHHKWNAADVVCQKCGCTALHPAGLPHLQSGSDVWLVDGVCSGN
ESFLWDCHRSQVNFDFLHONDVSVICSDGADLELRADGNNCSGRVERVTHEQWMT
ICDQNKNEALVVCQKCGCPSPFISGVVNSPAVLRIPI
DGKAKTCFRSDAGVICSADLRLVNGHSPCIHREYVIGQEWETVCHDRWSTR
NAAVVCRLQGPSPFISGVVNSPAVLRIPI
HREDVILVCGDNLVGLVGSNRCGRLEVIYFQGRWGTVCDDGNSKAAVVCQSL
DCPSIIIGGLGNAGTGYKIWDVSCDGEDSLWSCRNSGWNDCSHSDGVGIC
SDASMEILRVGSSRCAGKVEVNOVGAVGILCANGWMTAEVVCRLQEGSARVS
RPHFTETLHILMSNGCTGEASLWDCIRWKKOTACHLNMEASLICSARQPRLV
GADMPCSRVEVKHADTRVSRDSDFSLHAANVLCRLNCGDAISLSVDFHFGKNGL
TWAQFQEGSETHALCPVIOHPEDTCHIHREVGVCVRYTDVRLVNGKSCQDQVE
INVLHWSGLDTHMDPDRVLCQLSCGTALSTTGGKIGERSVRVWGHFHLGN
ESLDNCOMTVLGPAPPCIHGNTVSVICSTLQPLFLPCLANVSDPYSVAVPEGALIC
LEDKRLRLVDSRCAGRVEIYHDGFNGTICDDGWDLSDAHVCQKLCGVAFNATVS
AHFGSGSPIDWLLDNCVGMESHLMQCPSRGNGHDCRHKEDAGVICSEFTALRYSE
TETSCACRLVEFYNTGWSVRRNITTAIAGVLCRLGCGENGVSILAPSKVCSGF
MWVDICQPKTHISLWQCLSAPEWRISSTAEETWCTEDRIVRGGTETGSRVEIW
HAGSNGTVCDSDNLAEAEVVCQKCGSALALADSFQGGTGIWDDMRCKNES
FLWDCRHPWSDCGHREKEDAGVRCSGQSLKSNASSGRLALISLISFGLLPLVLFIL
FLTWCRVQKQHLPLRVSTRRRGSLLENLFHEMETCLKREDPHGTRTSDDTPNHGCE
ASDTSLLGLVPASEATK"
26. .151
152. .469
/note="Region: scavenger receptor cysteine-rich domain
SRCR 1"
470. .790
/note="Region: scavenger receptor cysteine-rich domain
SRCR 2"
791. .1111
/note="Region: scavenger receptor cysteine-rich domain
SRCR 3"
1112. .1432
/note="Region: scavenger receptor cysteine-rich domain
SRCR 4"
1433. .1753
/note="Region: scavenger receptor cysteine-rich domain
SRCR 5"
1754. .2074
/note="Region: scavenger receptor cysteine-rich domain
SRCR 6"
2075. .2395
/note="Region: scavenger receptor cysteine-rich domain
SRCR 7"
2396. .2710
/note="Region: scavenger receptor cysteine-rich domain

Query Match 98.2%; Score 4543.4; DB 9; Length 4579;
Best Local Similarity 99.8%; Pred No. 0;
Matches 4550; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 21 ACTAGTATGATGCTGCCCTCAAACTCGGCGATATGATTTGGAGATGCTGCTGCA 80
DB 19 ACCATAATGATGCTGCCCTCAAACTCGGCGATATGATTTGGAGATGCTGCTGCA 78
QY 81 TCAGACCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 140
DB 79 TCAGACCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 138
QY 141 CAGCAGTTTTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGAGACGGTCCCTGCTC 200
DB 139 CAGCAGTTTTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGAGACGGTCCCTGCTC 198
QY 201 TGGACAGTGGAGTGAATTCACAGGACAGTGGGGAGTGTGTGATGATGGTGGAA 260
DB 199 TGGACAGTGGAGTGAATTCACAGGACAGTGGGGAGTGTGTGATGATGGTGGAA 258
QY 261 CACTACTGCTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 320
DB 259 CACTACTGCTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 318
QY 321 TCGTTTTGGACAAGCCGCTGACTAGACATGGAATAATTTGGCTTGATGATGTTTTCCCTGTTA 380
DB 319 TCGTTTTGGACAAGCCGCTGACTAGACATGGAATAATTTGGCTTGATGATGTTTTCCCTGTTA 378
QY 381 TGGAAATGAGTCACTCTCTGCGAATGTCACACCGGGAATGGGAGCAATCAACTGTTA 440
DB 379 TGGAAATGAGTCACTCTCTGCGAATGTCACACCGGGAATGGGAGCAATCAACTGTTA 438
QY 441 TCATGGAGAAGATGTTGGTGTCAACTGTTATGTTGGAAGCAATCTGGGTTTCAGGCTACT 500
DB 439 TCATGGAGAAGATGTTGGTGTCAACTGTTATGTTGGAAGCAATCTGGGTTTCAGGCTACT 498
QY 501 GGATGGAACAACACTCCTGTTTCAGGAGAGTGGAGGTGAAATTCGAAGAAAGTGGGGAC 560
DB 499 GGATGGAACAACACTCCTGTTTCAGGAGAGTGGAGGTGAAATTCGAAGAAAGTGGGGAC 558
QY 561 TATATGTGATGATGGTGAACACTTGAATCTGCTCCGCTGGTGTGCGAGCACTAGGATG 620
DB 559 TATATGTGATGATGGTGAACACTTGAATCTGCTCCGCTGGTGTGCGAGCACTAGGATG 618
QY 621 TCCATCTCTCTTTTATTTCTCTGGAGTGTGTAATAGCCCTGCTGCTATTCGCGCCCATTTG 680
|||||

misc_feature
2711. .3025
/note="Region: scavenger receptor cysteine-rich domain
SRCR 9"
3026. .3118
/note="Region: interdomain segment"
3119. .3433
/note="Region: scavenger receptor cysteine-rich domain
SRCR 10"
3434. .3754
/note="Region: scavenger receptor cysteine-rich domain
SRCR 11"
3755. .4063
/note="Region: scavenger receptor cysteine-rich domain
SRCR 12"
4064. .4093
/note="Region: connecting peptide"
4094. .4171
/note="Region: transmembrane domain"
4172. .4384
/note="Region: cytoplasmatic domain"
4388. .4579
polyA_signal
4456. .4461
polyA_signal
4550. .4555
BASE COUNT 1147 a 919 c 1343 g 1170 t
ORIGIN

sig_peptide
misc_feature

misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature

Db	619	TTCCATCTTCTTTTATTTCTTCTGGAGTTGTAAATAGCCCTGCTGTATTTGGCGCCCAATTG	678
Qy	681	GCTGGATGACATTTTATGCCAGGGAATGAGTTGGCACTTGGAAATGCGACATCTGTGG	740
Db	679	GCTGGATGACATTTTATGCCAGGGAATGAGTTGGCACTCTGGAAATGCCAGACATCTGG	738
Qy	741	ATGGGGAATCATGACTGCAGTCACAATGAGAGTGTACATTAACCTTGTATATGATAGTAG	800
Db	739	ATGGGGAATCATGACTGCAGTCACAATGAGAGTGTACATTAACCTTGTATGATAGTAG	798
Qy	801	TGATCTTGAACATAAGCGTTGTAGTGGAACTAACCGCTGTATGGGAGAGTAGAGCTGAA	860
Db	799	TGATCTTGAACATAAGCGTTGTAGTGGAACTAACCGCTGTATGGGAGAGTAGAGCTGAA	858
Qy	861	AATCCAAGGAAGGTGGGGACCGTATGCCACCATAAGTGGAAACAATGCTGCAGCTGATGT	920
Db	859	AATCCAAGGAAGGTGGGGACCGTATGCCACCATAAGTGGAAACAATGCTGCAGCTGATGT	918
Qy	921	CGTATGCAAGCAGTTGGGATGTGGAAACGCACTTCACTTCGCTGGCTTGCCTCAATTTGCA	980
Db	919	CGTATGCAAGCAGTTGGGATGTGGAAACGCACTTCACTTCGCTGGCTTGCCTCAATTTGCA	978
Qy	981	GTCAAGGCTGTGATTTGATGCTTGATGGTGTCTCCCTGCTCCGTAATGAATCTTTTCT	1040
Db	979	GTCAAGGCTGTGATTTGATGCTTGATGGTGTCTCCCTGCTCCGTAATGAATCTTTTCT	1038
Qy	1041	TTGGGACTGCAGACATTTCCGGAAACCGTCAAAATTTTGACTGTCTTCAATAAACAGATGTCT	1100
Db	1039	TTGGGACTGCAGACATTTCCGGAAACCGTCAAAATTTTGACTGTCTTCAATAAACAGATGTCT	1098
Qy	1101	TGTTGATCTGCTCAGATGGAGCAGATTTGGAAGTGGCACTAGCAGATGAAGTAACAATTG	1160
Db	1099	TGTTGATCTGCTCAGATGGAGCAGATTTGGAAGTGGCACTAGCAGATGAAGTAACAATTG	1158
Qy	1161	TTCAAGGAGATGAGCGTCAGAAATTCATGAACAGTGTGGACAATATGTGTACAGAACTG	1220
Db	1159	TTCAAGGAGATGAGCGTCAGAAATTCATGAACAGTGTGGACAATATGTGTACAGAACTG	1218
Qy	1221	GAAGAATGAACAAGCCCTTTGTGGTTTGTAAAGCAGCTAGGATGTCCGTTACAGCGTCTTGG	1280
Db	1219	GAAGAATGAACAAGCCCTTTGTGGTTTGTAAAGCAGCTAGGATGTCCGTTACAGCGTCTTGG	1278
Qy	1281	CAGTCGTGTCGCTAAACCTTAGTAATGAAGCTAGAGACATTTGGATTAACAGCATACTTG	1340
Db	1279	CAGTCGTGTCGCTAAACCTTAGTAATGAAGCTAGAGACATTTGGATTAACAGCATACTTG	1338
Qy	1341	CAGTCGGAATGAGTCAGCTCTCGGACGTGCACATATCATGGAAGAACGAAGCGAACATG	1400
Db	1339	CAGTCGGAATGAGTCAGCTCTCGGACGTGCACATATCATGGAAGAACGAAGCGAACATG	1398
Qy	1401	CTTCCGAAGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTTAAGGCT	1460
Db	1399	CTTCCGAAGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTTAAGGCT	1458
Qy	1461	TGTCGGGGCTCATAGCCCTGTTATGGGAGATTTGAGTGAAATACGAAGAGAGTGGGG	1520
Db	1459	TGTCGGGGCTCATAGCCCTGTTATGGGAGATTTGAGTGAAATACGAAGAGAGTGGGG	1518
Qy	1521	GACTGTGTGTCATGACAGATGGAGCACAGGAATCCAGCTGTGTGTGTAAACAATTTGGG	1580
Db	1519	GACTGTGTGTCATGACAGATGGAGCACAGGAATCCAGCTGTGTGTGTAAACAATTTGGG	1578
Qy	1581	ATGTGGAAGCCCTATGCATGTCTTTGGTATGACCTTATTTTAAAGAAGCATCAGGACCTAT	1640
Db	1579	ATGTGGAAGCCCTATGCATGTCTTTGGTATGACCTTATTTTAAAGAAGCATCAGGACCTAT	1638
Qy	1641	TTGGCTGGATGACGTTTCTTGGCATTTGGAAATGAGTCAAAATATCTGGGACTGTGAACACAG	1700
Db	1639	TTGGCTGGATGACGTTTCTTGGCATTTGGAAATGAGTCAAAATATCTGGGACTGTGAACACAG	1698
Qy	1701	TGATGGGGAACCATTAATTTGTACACAGAGAGATGTGATTTGAACCTCTCAGGTGA	1760
Db	1699	TGATGGGGAACCATTAATTTGTACACAGAGAGATGTGATTTGAACCTCTCAGGTGA	1758

QY	1761	TCACAATGGGGCCCTGAGGCTGGTGGCGGCGCAGCAACCGCTGCTCGGGAAGACTGGAGGT	1820
DB	1759	TCACAATGGGGCCCTGAGGCTGGTGGCGGCGCAGCAACCGCTGCTCGGGAAGACTGGAGGT	1818
QY	1821	GTACTTTCAAGACAGGTGGGGCACAGTGTGTGATGACGGCTGGACACAGTAAAGCTCGAGC	1830
DB	1819	GTACTTTCAAGACAGGTGGGGCACAGTGTGTGATGACGGCTGGACACAGTAAAGCTCGAGC	1828
QY	1881	TGTGGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATTTGGCATGGGTCTGGGAACGC	1940
DB	1879	TGTGGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATTTGGCATGGGTCTGGGAACGC	1938
QY	1941	TTCTACAGGATATGGAATAATTTGGCTCGATGATGATTTCTCTGTGATGGAGATCAGTCAGA	2000
DB	1939	TTCTACAGGATATGGAATAATTTGGCTCGATGATGATTTCTCTGTGATGGAGATCAGTCAGA	1998
QY	2001	TCCTCTGGTCATCCAGGACAGTGGGTGGGGAATAATGACTGCAGTCCAGTCAAGATGT	2060
DB	1999	TCCTCTGGTCATCCAGGACAGTGGGTGGGGAATAATGACTGCAGTCCAGTCAAGATGT	2058
QY	2061	TGGAGTGATCTCTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAG	2120
DB	2059	TGGAGTGATCTCTTCTGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAG	2118
QY	2121	GTGTGCTGGAAAGTTGAGGTGAATGCCAGGTGCCGTGGGAATTTGTGTGCTTAATGG	2180
DB	2119	GTGTGCTGGAAAGTTGAGGTGAATGCCAGGTGCCGTGGGAATTTGTGTGCTTAATGG	2178
QY	2181	CTGGGGAATGACATTGCTGAAGTTGTTTTCGAGGCAACTTGAATGTGGGTCTGCAATCAG	2240
DB	2179	CTGGGGAATGACATTGCTGAAGTTGTTTTCGAGGCAACTTGAATGTGGGTCTGCAATCAG	2238
QY	2241	GGTCTCCAGAGAGCCCTCATTTTCACAGAAAGACATTTACACATCTTAATGTCGAATTTCTGG	2300
DB	2239	GGTCTCCAGAGAGCCCTCATTTTCACAGAAAGACATTTACACATCTTAATGTCGAATTTCTGG	2298
QY	2301	CTGCACTGGAGGGGAAGCCCTCTCTCTGGGATTTATACGATGGGATGGAAACAGACTGC	2360
DB	2299	CTGCACTGGAGGGGAAGCCCTCTCTCTGGGATTTATACGATGGGATGGAAACAGACTGC	2358
QY	2361	CTGTCTCAATTTAAATATGGAAGCAGTTTGATCTCTCAGCCACAGGCGCCAGGCTGGT	2420
DB	2359	CTGTCTCAATTTAAATATGGAAGCAGTTTGATCTCTCAGCCACAGGCGCCAGGCTGGT	2418
QY	2421	TGGAGCTGATATGCCCTGCTCTGGAGCTGTTGAAGTGAACATGCGACACATGCGCGCTC	2480
DB	2419	TGGAGCTGATATGCCCTGCTCTGGAGCTGTTGAAGTGAACATGCGACACATGCGCGCTC	2478
QY	2481	TGCTGTGATATCTGATTTCTCTCTTCATGTCGCCAATGTGCTGTGCAGAGAAATTAATTTG	2540
DB	2479	TGCTGTGATATCTGATTTCTCTCTTCATGTCGCCAATGTGCTGTGCAGAGAAATTAACNTG	2538
QY	2541	TGGAGATGCCATATCTTTCTGTGGGAGATCACTTTGAAAGGGAATTTGGTCTAACTTG	2600
DB	2539	TGGAGATGCCATATCTTTCTGTGGGAGATCACTTTGAAAGGGAATTTGGTCTAACTTG	2598
QY	2601	GGCCGAAAAAGTTCCAGTGTGAAGGGAGTGAACCTCACTTGCATTTATGCCCCCATTTGTCA	2660
DB	2599	GGCCGAAAAAGTTCCAGTGTGAAGGGAGTGAACCTCACTTGCATTTATGCCCCCATTTGTCA	2658
QY	2661	ACATCCGGAAGACATTTGATCCACAGCAGAGAAAGTTGGAGTTGTCTGTTCGCCGATATAC	2720
DB	2659	ACATCCGGAAGACATTTGATCCACAGCAGAGAAAGTTGGAGTTGTCTGTTCGCCGATATAC	2718
QY	2721	AGATGTCGCACTTTGTAATGGCAATCCCACTGTGACGGCGCAAGTGGAGTCAACGTGCT	2780
DB	2719	AGATGTCGCACTTTGTAATGGCAATCCCACTGTGACGGCGCAAGTGGAGTCAACGTGCT	2778
QY	2781	TGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCAGAGATGCCGTGTTCTATG	2840
DB	2779	TGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCAGAGATGCCGTGTTCTATG	2838

QY 2841 CAGACAGCTCAGCTGTGGGAGCTCTCTCTCAACACAGAGGAGAAATATATGGAGAAAG 2900
DB 2839 CAGACAGCTCAGCTGTGGGAGCTCTCTCTCAACACAGAGGAGAAATATATGGAGAAAG 2898
QY 2901 AAGTGTTCGTGTGGGAGACAGAGTTTCATTCCTTAGGAATGAGTCACTCTGGATAA 2960
DB 2899 AAGTGTTCGTGTGGGAGACAGAGTTTCATTCCTTAGGAATGAGTCACTCTGGATAA 2958
QY 2961 CTGTCAANTACAGAGTCTTGGAGACCTCCCTGTATCCATGGAATACATCTCTGTGAT 3020
DB 2959 CTGTCAANTACAGAGTCTTGGAGACCTCCCTGTATCCATGGAATACATCTCTGTGAT 3018
QY 3021 CTCACAGGAAGCCTGACCCAGACCTGTTTCCATCCCTCGCAAAATGTATCTCAACCATTA 3080
DB 3019 CTCACAGGAAGCCTGACCCAGACCTGTTTCCATCCCTCGCAAAATGTATCTCAACCATTA 3078
QY 3081 TTTGTCTGAGTTCACAGAGGCGAGTCTTTGTATCTCTTTAGGAGACAAAGCTCTCGCCT 3140
DB 3079 TTTGTCTGAGTTCACAGAGGCGAGTCTTTGTATCTCTTTAGGAGACAAAGCTCTCGCCT 3138
QY 3141 AGTGGATGGGACAGCGCTGTCCGGGAGAGTAGAGATCTATCAGACGGCTTCTGGGG 3200
DB 3139 AGTGGATGGGACAGCGCTGTCCGGGAGAGTAGAGATCTATCAGACGGCTTCTGGGG 3198
QY 3201 CACCATCTGTGATCAGCGCTGGGACCTGTAGCGATGCCAGTGGTGTCGCAAGCTGGG 3260
DB 3199 CACCATCTGTGATCAGCGCTGGGACCTGTAGCGATGCCAGTGGTGTCGCAAGCTGGG 3258
QY 3261 CTGTGGAGTGGCTTCAATGCCAGCGTCTCTGCTCACTTTGGGAGGGGTGACGGCCAT 3320
DB 3259 CTGTGGAGTGGCTTCAATGCCAGCGTCTCTGCTCACTTTGGGAGGGGTGACGGCCAT 3318
QY 3321 CTGCTCGATGACCTGAACCTGCACAGGAACGGAGTCCCACTTGTGGCAGTCCCTTCCCG 3380
DB 3319 CTGCTCGATGACCTGAACCTGCACAGGAATGGAGTCCCACTTGTGGCAGTCCCTTCCCG 3378
QY 3381 CGGCTGGGGGACGACGACCTGCAGGACACAGAGGAGCGAGGGTCACTCTCTCAAAAT 3440
DB 3379 CGGCTGGGGGACGACGACCTGCAGGACACAGAGGAGCGAGGGTCACTCTCTCAAAAT 3438
QY 3441 CACAGCCTTGAGGCTTACAGTGAACCTGAACAGAGAGCTGTCTCGGAGATGGAGT 3500
DB 3439 CACAGCCTTGAGGCTTACAGTGAACCTGAACAGAGAGCTGTCTCGGAGATGGAGT 3498
QY 3501 CTTCTATAACGGGACCTGGGACGCTCGGACGAGGAACATCACCAAGACATAGCAGG 3560
DB 3499 CTTCTATAACGGGACCTGGGACGCTCGGACGAGGAACATCACCAAGACATAGCAGG 3558
QY 3561 CATTGTGTGACGAGCTGGGCTGTGGGAGAAATGGAGTTGTACGCTCTCTTATC 3620
DB 3559 CATTGTGTGACGAGCTGGGCTGTGGGAGAAATGGAGTTGTACGCTCTCTTATC 3618
QY 3621 TAAGACAGGCTCTGTTTCATGTGGTGGAGACATTCAGTGTCTTAAACGCAATCTC 3680
DB 3619 TAAGACAGGCTCTGTTTCATGTGGTGGAGACATTCAGTGTCTTAAACGCAATCTC 3678
QY 3681 CATATGGCAGTGCCTGTCTGCCCATGGGACGAGGAATCTCCAGCCACGAGAGAGAC 3740
DB 3679 CATATGGCAGTGCCTGTCTGCCCATGGGACGAGGAATCTCCAGCAGCAGAGAGAC 3738
QY 3741 CTGGATACATGTGAAGATGAATAAGAGTGGTGGAGGACACCGAGTCTCTGGAG 3800
DB 3739 CTGGATACATGTGAAGATGAATAAGAGTGGTGGAGGACACCGAGTCTCTGGAG 3798
QY 3801 AGTGGAGATCTGGCAGCAGCTCTCTGGGACACAGTGTGATGACTCTCTGGGACCTGGC 3860
DB 3799 AGTGGAGATCTGGCAGCAGCTCTCTGGGACACAGTGTGATGACTCTCTGGGACCTGGC 3858
QY 3861 CGAGCGGAAGTGTGTCTCAGCAGCTGGGCTGTGGCTCTCTCTGGCTGCCCTGAGGGA 3920
DB 3859 CGAGCGGAAGTGTGTCTCAGCAGCTGGGCTGTGGCTCTCTCTGGCTGCCCTGAGGGA 3918
QY 3921 CGCTTCTGTTGGCCAGGGAACCTGGAACCATCTGGTTGGATGACATGCGGTCCAAAGGAA 3980

DB 3919 CGCTTCGTTTGGCCAGGGAACCTGGAAACCTCTGGTTGGATGACATGCGGTGCAAGGAA 3978
QY 3981 TGAGTCAATTTCTATGGGACTGTCAAGCCCAACCCCTGGGACAGAGTGAATGTGGACAA 4040
DB 3979 TGAGTCAATTTCTATGGGACTGTCAAGCCCAACCCCTGGGACAGAGTGAATGTGGACAA 4038
QY 4041 GGAAGATGCTGGCGTGGAGTGTCTGGAGACGTCGCTGAATCACTGAATGCTCTCTCAGG 4100
DB 4039 GGAAGATGCTGGCGTGGAGTGTCTGGAGACGTCGCTGAATCACTGAATGCTCTCTCAGG 4098
QY 4101 TCATTTAGCACTATTTATCCAGTATCTTTGGGCTCTCTTCCTCGGTTCTCTTTATCT 4160
DB 4099 TCGTTTAGCACTATTTATCCAGTATCTTTGGGCTCTCTTCCTCCGGTCTCTTTATCT 4158
QY 4161 ATTTCTCAGTGTGCGGAGTTTCAGAAACAAAACATCTGCCCTCAGAGTTTCAACGAG 4220
DB 4159 ATTTCTCAGTGTGCGGAGTTTCAGAAACAAAACATCTGCCCTCAGAGTTTCAACGAG 4218
QY 4221 AAGGAGGGTCTCTCGAGGAGAAATTTATTCATGAGATGGAGACCTGCCTCAAGAGAG 4280
DB 4219 AAGGAGGGTCTCTCGAGGAGAAATTTATTCATGAGATGGAGACCTGCCTCAAGAGAG 4278
QY 4281 GGACCCACATGGACAAAGAACCTCAGATGACACCCCAACCATGTTGTGAAGATGCTAG 4340
DB 4279 GGACCCACATGGACAAAGAACCTCAGATGACACCCCAACCATGTTGTGAAGATGCTAG 4338
QY 4341 CGACACATCGCTGTGGGAGTCTTCTCCTCGCTCTGAAGCCCAAAATGACTTTAGACTTC 4400
DB 4339 CGACACATCGCTGTGGGAGTCTTCTCCTCGCTCTGAAGCCCAAAATGACTTTAGACTTC 4398
QY 4401 CAGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAGACAAACACTTTTAAATGAAT 4460
DB 4399 CAGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAGACAAACACTTTTAAATGAAT 4458
QY 4461 AAAGAGCAAGTCAAGTTGCCCTATGAAAACCTTGTCCAAATACACTTTTGAACAAATAG 4520
DB 4459 AAAGAGCAAGTCAAGTTGCCCTATGAAAACCTTGTCCAAATACACTTTTGAACAAATAG 4518
QY 4521 GAGAAGCTTAAATGATAAGACTGGTGAATAATAAAATTTGAATTTATGATACACTGT 4580
DB 4519 GAGAAGCTTAAATGATAAGACTGGTGAATAATAAAATTTGAATTTATGATACACTGT 4578
QY 4581 T 4581
DB 4579 T 4579

RESULT 2

AX133829

LOCUS

Sequence 15 from Patent WO0119856.

DEFINITION

AX133829

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

AX133829 3670 bp DNA linear PAT 15-MAY-2001
Sequence 15 from Patent WO0119856.
AX133829
AX133829.1 GI:14139781

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 3670)

Shimkets, R.A., Fernandes, E., Herrmann, J.L., Liu, X., Yang, M. and

Beldog, F.L.

Secreted human proteins, polynucleotides encoding them and methods

of using the same

Patent: WO 0119856-A 15 22-MAR-2001;

Curagen Corporation (US)

Location/Qualifiers

1..3670

/organism="Homo sapiens"

/db_xref="taxon:9606"

113..3475

/note="unnamed protein product"

/codon_start=1

/protein_id="CAC38965.1"
/db_xref="gi:14139782"
/translation="MKNDADVKKIRLENPNSENKVPTRHKMDGADLELRADGNSNN
ISGVEVRIHEDWTTICDONKNEQALVCKQGCPEFSGSRRAKPSNEARDLWINS
CSTGNEALHEDVYDCKAKKTCFRSDAGVIGCDKADLRLVGAHSPCVRLEVKY
QEWGTVCHDRWRNARNAVVCCKQGCGRPHVFGMTYFKEASGPIWLDVDSICIGNESN
IWDCEHSGMVKHNCVHREDVITVTCSDGATWGLRVGNSRCSGRLEVYFQGRWGTVCSD
DWNKAAAVVCSQDLCPSIIIGMLGNASTGYGKIWLDVDSYCDGSDSLWSCNRSWG
GNDCSHSDGVYICSDASOMELRVGSSRCAGKVEVNOGAVGILICANGWOTACHLNAE
VNDQECGSAIVYSREPHFTERTLHILMSNSGCTGGEGEASLWDCIRWENKOTACHLNAE
EASLISAHROPRLVGLADPCSGRVEVKHADTWRSYCDSESLHAANVLCRELNGDA
ISLSVDFHFGKGNELWAEKFCQEGSETHLALCPVQHPEDTCHSREVGLVGSRYTD
ILVYNGSCQDQVEIINWGHWSLCTDHPEDARVLCRLSCGTALVSTGGYIGY
RSVWRHREHCLNESLNDQNTVLGAPPCIHGNTVSVICTSLQPLPCLANVS
DPLYSAYVEFSALICLEDKRLRVGDSRCAGRVEIYHDFEWGTICDDGMDLSLNAVY
CKLGGVAFNAPVSAHFEGEGPIWLDLNLCTGMSHLCWCPKSRGQDCHRIKEDA
GVICSEETALRYSETETGSCAGLEVYNGTWGSGVRNITTAAGLVCHRCOLCGEN
GVSLAPLSKTSQGFPMVDDIQCPKTHISIMQCLSAWPKRISPAEETWICEDRIR
VRGDETCGRVEIWHAGSMGTVCDDSDWDLAEVVCQGLCGSALALRSDASFGQT
GTIWLDDMRCKGNESFLWDCHEKAPWQSDGCKEDAGVCSQSLKSLNASGSHLALI
LSSIFGLILLVLFILFTWCRVQKHLPLRVSTRRSGSLEENLFHEMETCLKREDPH
GTRTSDDTPNHGCEADSDTSLGLVLPASEATK"

BASE COUNT 947 a 752 c 1073 g 898 t

ORIGIN

Query Match 74.8%; Score 3459.8; DB 6; Length 3670;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3464; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY	1174	CTGAGGCTGGTGGGGCGGAGCAACCGCTGCTCGGGAAGACTGAGGAGTGTACTTTCAAGGA	1833
DB	860	CTGAGGCTGGTGGGGCGGAGCAACCGCTGCTCGGGAAGACTGAGGAGTGTACTTTCAAGGA	919
QY	1834	CGGTGGGCGACAGTGTGTGATGACGGCTGGACACATAAGCTGCACCTGTGGTGTGTAGC	1893
DB	920	CGGTGGGCGACAGTGTGTGATGACGGCTGGACACATAAGCTGCACCTGTGGTGTGTAGC	979
QY	1894	CAGCTGGAGCTGCCCATCTTCTATATTGCTGGGTCTGGGAAACGCTTCTAGAGGATAT	1953
DB	980	CAGCTGGAGCTGCCCATCTTCTATATTGCTGGGTCTGGGAAACGCTTCTAGAGGATAT	1039
QY	1954	GGAAAAATTTGGCTCGATGATCTTCTCTGTGATGAGATGAGTCAAGTCTCTGGTCTATGC	2013
DB	1040	GGAAAAATTTGGCTCGATGATCTTCTCTGTGATGAGATGAGTCAAGTCTCTGGTCTATGC	1099
QY	2014	AGGAACAGCTGGGTGGGGAATTAATGACTGCAGTCAAGTCAAGTCAAGTCAAGTCAAGT	2073
DB	1100	AGGAACAGCTGGGTGGGGAATTAATGACTGCAGTCAAGTCAAGTCAAGTCAAGTCAAGT	1159
QY	2074	TCTGATGCAATCGATATGAGCTGAGGCTTGTGGGTGGGAAGCAGCAGCTGCTGCAAAA	2133
DB	1160	TCTGATGCAATCGATATGAGCTGAGGCTTGTGGGTGGGAAGCAGCAGCTGCTGCAAAA	1219
QY	2134	GTGAGGTGAATGCTCCAGGCTCCGCTGGGAATTCCTGTGCTAATGGCTGGGGAATGAAC	2193
DB	1220	GTGAGGTGAATGCTCCAGGCTCCGCTGGGAATTCCTGTGCTAATGGCTGGGGAATGAAC	1279
QY	2194	ATTGCTGAAGTTGTTGACAGGCAACTTGAATGTGGGTCTGCAATCAGGCTCTCCAGAGAG	2253
DB	1280	ATTGCTGAAGTTGTTGACAGGCAACTTGAATGTGGGTCTGCAATCAGGCTCTCCAGAGAG	1339
QY	2254	CCTCATTTTACAGAAAGAACATTACACATCTTAATGCTCAATTTCTGGCTGCACTGGAGG	2313
DB	1340	CCTCATTTTACAGAAAGAACATTACACATCTTAATGCTCAATTTCTGGCTGCACTGGAGG	1399
QY	2314	GAAGCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACACTGCGTGTCAATTAAT	2373
DB	1400	GAAGCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACACTGCGTGTCAATTAAT	1459
QY	2374	ATGGAAGCAAGTTTGTATCTGCTCAGCCACAGCCAGCCAGGCTGTTGGAGCTGTATATG	2433
DB	1460	ATGGAAGCAAGTTTGTATCTGCTCAGCCACAGCCAGCCAGGCTGTTGGAGCTGTATATG	1519
QY	2434	CCCTGCTCTGAGAGTGTGTAAGTGAACATGAGACACATGCGCTCTCTGTGTGTCT	2493
DB	1520	CCCTGCTCTGAGAGTGTGTAAGTGAACATGAGACACATGCGCTCTCTGTGTGTCT	1579
QY	2494	GATTTCTCTCTCTGCTGCCAATGTGCTGTCAGAGAAATTAATTTGTCAGATGTCATAT	2553
DB	1580	GATTTCTCTCTCTGCTGCCAATGTGCTGTCAGAGAAATTAATTTGTCAGATGTCATAT	1639
QY	2554	TCTCTTTCTGTCGGGAGATCATTGGAAGGAATGGTCTAATTTGGGCCGCAAAAGTTC	2613
DB	1640	TCTCTTTCTGTCGGGAGATCATTGGAAGGAATGGTCTAATTTGGGCCGCAAAAGTTC	1699
QY	2614	CAGCTGGAAGGAGTCAAACTCACCTTGCATTTGCCCATTTGTTCAACATCCGGAAGAC	2673
DB	1700	CAGCTGGAAGGAGTCAAACTCACCTTGCATTTGCCCATTTGTTCAACATCCGGAAGAC	1759
QY	2674	ACTTCTATCCACAGAGAGAACTTGGAGTTGCTGTCTCCCGATATACAGATGTCGACAT	2733
DB	1760	ACTTCTATCCACAGAGAGAACTTGGAGTTGCTGTCTCCCGATATACAGATGTCGACAT	1819
QY	2734	GTGAATGGCAATCCAGTGTGACGGGCAAGTGGAGATCAACGTCTTGGACACTGGGGC	2793
DB	1820	GTGAATGGCAATCCAGTGTGACGGGCAAGTGGAGATCAACGTCTTGGACACTGGGGC	1879
QY	2794	TCACCTGTGTACACCACTGGGACCCAGAGATGCCGTGTTCTATGACAGAGCTCAGC	2853
DB	1880	TCACCTGTGTACACCACTGGGACCCAGAGATGCCGTGTTCTATGACAGAGCTCAGC	1939
QY	2854	TCTGGGAGTCTCTCTCAACCACAGGAGGAAATATATTGGAGAAAGAGTGTCTGTGTG	2913

Db	3020	CAGGAACTCGAACCATCTGGTTGGATGACATCGCGTGCAAGAGAAATGAGTCATTTCTA	3079
Qy	3994	TGGGACTGTGCACGCCAAACCCCTGGGACAGAGTGACTGTGGACACAAGAGATGCTGGC	4053
Db	3080	TGGGACTGTGCACGCCAAACCCCTGGGACAGAGTGACTGTGGACACAAGAGATGCTGGC	3139
Qy	4054	GTGAGGTGCTCTGGACAGTCGCTGGAATCAGTCGAATGCCTCTCAGGTCATTTAGCACTT	4113
Db	3140	GTGAGGTGCTCTGGACAGTCGCTGGAATCAGTCGAATGCCTCTCAGGTCATTTAGCACTT	3199
Qy	4114	ATTTTATCCAGTATCTTTTGGGCTCCCTCTCCCTGGTTCTCTTTATTCATTTCTCACGTGG	4173
Db	3200	ATTTTATCCAGTATCTTTTGGGCTCCCTCTCCCTGGTTCTCTTTATTCATTTCTCACGTGG	3259
Qy	4174	TGCGAGTTCAGAAACAAAACATCTGCCCTCAGAGTTTCAACAGAGGAGGGTGTCT	4233
Db	3260	TGCGAGTTCAGAAACAAAACATCTGCCCTCAGAGTTTCAACAGAGGAGGGTGTCT	3319
Qy	4234	CTCGAGGAGATTTTATTCATGAGATGGAGACCTGCCTCAAGAGAGAGACCCACATGGG	4293
Db	3320	CTCGAGGAGATTTTATTCATGAGATGGAGACCTGCCTCAAGAGAGAGACCCACATGGG	3379
Qy	4294	ACAAGAACCTCAGATGACACCCCAACCATGGTTGTGAAGATGCTAGCGACACATCGGTG	4353
Db	3380	ACAAGAACCTCAGATGACACCCCAACCATGGTTGTGAAGATGCTAGCGACACATCGGTG	3438
Qy	4354	TTGGGAGTTCTTCTCGCTCTGAAGCCACAAAATGACTTTTAGACTTCCAGGCTCACCAG	4413
Db	3440	TTGGGAGTTCTTCTCGCTCTGAAGCCACAAAATGACTTTTAGACTTCCAGGCTCACCAG	3499
Qy	4414	ATCAACCTCTAAATATCTTTGAAGGAGACACAACTTTTAAATGATAAGAGGAGTGCA	4473
Db	3500	CTCAACCTCTAAATATCTTTGAAGGAGACAACTTTTAAATGATAAGAGGAGTGCA	3559
Qy	4474	AGTTTGCCTTATGGAAACTTTGTCCCAATACATTTCTTGAACAATAAGGAAACAGCTAAA	4533
Db	3560	AGTTTGCCTTATGGAAACTTTGTCCCAATACATTTCTTGAACAATAAGGAAACAGCTAAA	3619
Qy	4534	TTGATAAGACTGGTGATAATAAAAAATTTGAATTTATCTATATCAGCTGTTAAA	4584
Db	3620	TTGATAAGACTGGTGATAATAAAAAATTTGAATTTATCTATATCAGCTGTTAAA	3700

AK058038	AK058038	2249 bp	mrna	linear	PRI 31-OCT-2001
LOCUS	Homo sapiens	cDNA FLJ25309 fis, clone SRN00987			
DEFINITION	Homo sapiens	scavenger receptor cysteine-rich type 1 protein, Mr60 precursor.			

VERSION AK058038.1 GI:16554040

SOURCE
HOMO SAPIENS SYNOVIAL MEMBRANE (KNEE) CDN
clone: SYN00985

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Placentalia; Artiodactyla; Cetartiodactyla; Cetacea; Mysticeti; Balaenidae; *Balaenoptera*; *B. physalus*

REFERENCES
AUTHORS
Ishibashi, T., Kanehori, K., Yosida, M., Wata-

Kusano, J., Watanabe, M., Fujimori, K., Tanaka, Y., Chiba, Y., Suzuki, Y., Hata, H., Yamachita, Y.

Nishikawa, T., Sugiyama, A., Kawakami, B., Na

FILE	NEDO human cDNA sequencing project
JOURNAL	Unpublished

AUTHORS
Sugano, S. and Suzuki, Y.
Dietrich, C. H. M. J. J.

University of Tokyo, Laboratory of Genome

(E-mail: cdnale@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,
Fax: 81-3-5449-5416)

COMMENT
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan: cDNA full insert sequencing:
Research Association for Biotechnology (RAB): cDNA library
construction and 5'-end one pass sequencing: Institute of Medical
Science, University of Tokyo, Laboratory of Genome Structure, Human
Genome Center; 3'-end one pass sequencing: RAB; clone selection for
full insert sequencing: RAB and Helix Research Institute.

FEATURES
Location/Qualifiers

1..2249

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="SYN00985"

/tissue_type="synovial membrane (knee)"

/clone_lib="SVN"

/note="cloning vector: pME18SPL3"

559 a 512 c 638 g 540 t

BASE COUNT
ORIGIN

Query Match 48.5%; Score 2244.2; DB 9; Length 2249;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2246; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2333 GTATACGATGGAGTGGAAACAGACTGCGTGTCTCAATTTAAATATGGAAGCAAGTTGATCT 2392

Db 1 GTATACGATGGAGTGGAAACAGACTGCGTGTCTCAATTTAAATATGGAAGCAAGTTGATCT 60

Qy 2393 GCTCAGCCACAGGAGCCAGGCTGGTTGGAGCTGATATGCCCTGCTCTGGAGCTGTTG 2452

Db 61 GCTCAGCCACAGGAGCCAGGCTGGTTGGAGCTGATATGCCCTGCTCTGGAGCTGTTG 120

Qy 2453 AAGTGAACATCGACACATGCGCTCTGCTGTGATTCGATTCCTCTCTTCATGCTG 2512

Db 121 AAGTGAACATCGACACATGCGCTCTGCTGTGATTCGATTCCTCTCTTCATGCTG 180

Qy 2513 CCAATGTGCTGTCAGAGAAATTAATTTGTGGAGATGCCATATCTCTTTCTGTGGAGATC 2572

Db 181 CCAATGTGCTGTCAGAGAAATTAATTTGTGGAGATGCCATATCTCTTTCTGTGGAGATC 240

Qy 2573 ACTTTGGAAGGNAAGTCTTAACCTTGGGCCGAAAGTTCCAGTGTGAAGGGAGTGAAA 2632

Db 241 ACTTTGGAAGGNAAGTCTTAACCTTGGGCCGAAAGTTCCAGTGTGAAGGGAGTGAAA 300

Qy 2633 CTCACCTTGCAATATGCCCCAATGTTCAACATCCGGAAGACATTTGATCCACAGCAGAG 2692

Db 301 CTCACCTTGCAATATGCCCCAATGTTCAACATCCGGAAGACATTTGATCCACAGCAGAG 360

Qy 2693 AAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCGGACTTTGTAATGGCAAAATCCCACT 2752

Db 361 AAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCGGACTTTGTAATGGCAAAATCCCACT 420

Qy 2753 GTGAGGGCAAGTGGAGATCAACGTGCTTGACACTGGGGCTCACTGTGTACACCCCACT 2812

Db 421 GTGAGGGCAAGTGGAGATCAACGTGCTTGACACTGGGGCTCACTGTGTACACCCCACT 480

Qy 2813 GGGACCCAGAGATGCCCGTGTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAA 2872

Db 481 GGGACCCAGAGATGCCCGTGTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAA 540

Qy 2873 CCACAGGAGAAATATATTTGGAGAAAGTGTGCTGTGGGGAGACAGGTTTCATT 2932

Db 541 CCACAGGAGAAATATATTTGGAGAAAGTGTGCTGTGGGGAGACAGGTTTCATT 600

Qy 2933 GCTTAGGAATGAGTCACCTTCTGTAATCTCAATGACAGTTCCTGGAGCACCCTCCCT 2992

Db 601 GCTTAGGAATGAGTCACCTTCTGTAATCTCAATGACAGTTCCTGGAGCACCCTCCCT 660

Qy 2993 GTATCCATGGAATACTGCTCTGTGATCTGCACAGGAAGCTGACCCAGCCACTGTTTC 3052

Db 661 GTATCCATGGAATACTGCTCTGTGATCTGCACAGGAAGCTGACCCAGCCACTGTTTC 720

Qy 3053 CATGCCCTCGCAATGTATCTGACCCCATATTTGCTGCAAGTTCCAGAGGGGAGTGTCTTGA 3112

Db 721 CATGCTCGCAAAATGTATCTGACCCATATTTGCTCTGCAAGTTCCAGAGGGCAGTCTTGA 780

Qy 3113 TCTCTTTAGAGACAAACGGCTCCGCTTAGTGGATGGGACAGCCGCTCTCCCGGAGAG 3172

Db 781 TCTCTTTAGAGACAAACGGCTCCGCTTAGTGGATGGGACAGCCGCTCTCCCGGAGAG 840

Qy 3173 TAGAGATCTATACAGACAGCGCTTCTGGGGCAGCATCTGTGATGACGCTGGGACCTGAGCG 3232

Db 841 TAGAGATCTATACAGACAGCGCTTCTGGGGCAGCATCTGTGATGACGCTGGGACCTGAGCG 900

Qy 3233 ATGCCACGTTGGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCAGGCTCTCTG 3292

Db 901 ATGCCACGTTGGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCAGGCTCTCTG 960

Qy 3293 CTCACCTTTGGGAGGGGTTCAGGCCCCATCTGGCTGGATGACCTGAACCTGCACAGAAATGG 3352

Db 961 CTCACCTTTGGGAGGGGTTCAGGCCCCATCTGGCTGGATGACCTGAACCTGCACAGAAATGG 1020

Qy 3353 AGTCCCACTTGTGCAAGTGCCTTCCCGCGCTGGGGCAGCAGCAGCTGCAGGCAACAAG 3412

Db 1021 AGTCCCACTTGTGCAAGTGCCTTCCCGCGCTGGGGCAGCAGCAGCTGCAGGCAACAAG 1080

Qy 3413 AGGAGCGAGGGTCACTGTCTCAGAAATTCACAGCCTTGAGGCTCTACAGTGAACCTGAAA 3472

Db 1081 AGGAGCGAGGGTCACTGTCTCAGAAATTCACAGCCTTGAGGCTCTACAGTGAACCTGAAA 1140

Qy 3473 CAGAGAGCTGTCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGCGACGCTCGGCA 3532

Db 1141 CAGAGAGCTGTCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGCGACGCTCGGCA 1200

Qy 3533 GGAGGAACATCACCACAGCCATAGCAGCATTTGTGTGAGGAGCTGGGCTGTGGGGAGA 3592

Db 1201 GGAGGAACATCACCACAGCCATAGCAGCATTTGTGTGAGGAGCTGGGCTGTGGGGAGA 1260

Qy 3593 ATGGAGTTGTGAGGCTCGCCCTTTATCTAAGACAGGCTCTGTTTTCATGTGGGTGGATG 3652

Db 1261 ATGGAGTTGTGAGGCTCGCCCTTTATCTAAGACAGGCTCTGTTTTCATGTGGGTGGATG 1320

Qy 3653 ACATTCAGTGTCTTAAACGCATATCTCCATATGCACTGCTCTGCTCCCATCGGAGC 3712

Db 1321 ACATTCAGTGTCTTAAACGCATATCTCCATATGCACTGCTCTGCTCCCATCGGAGC 1380

Qy 3713 CAAGAATCTCCAGCCAGCAGAGAGCTGATCACATGTGAAGATAGAATAAGAGTGC 3772

Db 1381 CAAGAATCTCCAGCCAGCAGAGAGAGCTGATCACATGTGAAGATAGAATAAGAGTGC 1440

Qy 3773 GTGAGGAGACACCGAGTGTCTTGGGAGTGGAGATCTGGCACCGCAGGCTCTCTGGGCA 3832

Db 1441 GTGAGGAGACACCGAGTGTCTTGGGAGTGGAGATCTGGCACCGCAGGCTCTCTGGGCA 1500

Qy 3833 CAGTGTGTGATGACTCTCTGGGACCTGGCGGAGGAGTGTGTGTGAGCAGCTGGGCT 3892

Db 1501 CAGTGTGTGATGACTCTCTGGGACCTGGCGGAGGAGTGTGTGTGAGCAGCTGGGCT 1560

Qy 3893 GTGGCTGTGCTCTGCTGCTGCTGAGGAGCGCTTCTTTGGCCAGGGAACCTGGAACCATCT 3952

Db 1561 GTGGCTGTGCTCTGCTGCTGAGGAGCGCTTCTTTGGCCAGGGAACCTGGAACCATCT 1620

Qy 3953 GTTGTGGATGACATGCGGTGTCAAAAGAAATGAGTCAATTTCTATGGAGCTGTACGCCAAG 4012

Db 1621 GTTGTGGATGACATGCGGTGTCAAAAGAAATGAGTCAATTTCTATGGAGCTGTACGCCAAG 1680

Qy 4013 CCTGGGACAGAGTCACTGTGGACACAAGGAAGTGTGGGCTGAGGTGCTCTGGACAGT 4072

Db 1681 CCTGGGACAGAGTCACTGTGGACACAAGGAAGTGTGGGCTGAGGTGCTCTGGACAGT 1740

Qy 4073 CGCTGAATCACTCAATGCCCTCCTCAGCTCAATTTAGCACTTATTTATCCAGTATCTTTG 4132

Db 1741 CGCTGAATCACTCAATGCCCTCCTCAGCTCAATTTAGCACTTATTTATCCAGTATCTTTG 1800

Qy 4133 GGCTCCCTCTCTGCTGTTTATTTCTTATTTCTCACGTGTGGCGAGTTCACAAACAA 4192

Qy 1659 TTGCATTGGAATGAGTCAAAATATCTGGACTGTGAACACAGTGTGATGGGAAAGCATAA 1718
Dy 1767 ATGCAACGGAATGAGTCAGCTCTGGAACATGCAAAACATCAAGATGGGAAAGCATAA 826
Qy 1719 TTGTGTACACAGAGAGATGTGATTAACCTGCTCAGGTGATGCAACATGGGCGCTGAG 1778
Dy 827 CTGTGATCATCTGAGGATGCTGAGTGTATTTGCTCAAGGAGGACAGATCTGAGCCTGAG 886
Qy 1779 GCTGTGGGCGGACGACACCCCTGCTCGGGAAGAGTGGAGGTGTACTTTCAAGGAGCGTG 1838
Dy 887 ACTGGTAGATGGATCACTGAATGTTACGGAAGATAGAGTGAGATTTCCAAGGAGATG 946
Qy 1839 GGGCACAGTGTGTGATGACGGCTGGACAGATAAAGCTGCAGCTGTGTGTAGCCAGCT 1898
Dy 947 GGGGACAAATATGTGATGACGGCTGGGACAGTTACGATGCTGTGTGGCATSCAAGCACT 1006
Qy 1899 GGACTGCCCATCTTCATCATTTGATGGATGGGTCTGGGAACCCCTCTACAGGATATGGAAA 1958
Dy 1007 GGGATGTCCAACTCCGCTCACAGCCATTTGGTCGAGTTAAACGCCAGTAAGGGAITTTGGACA 1066
Qy 1959 AATTGGCTCGATGATGTTTCTGTGTGATGGAGATGAGTCAGATCTCTGCTCATGCAAGAA 2018
Dy 1067 CATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTGTGGCAATGTAACA 1126
Qy 2019 CAGTGGGTGGGAAATATGACTGCAGTCACAGTGAAGATTTGGAGTGATCTGTCTGA 2078
Dy 1127 CCATGAATGGGAAAGCATTATTGCAATCACAATGAAGATGCTGGCGTGACATGTTCTGA 1186
Qy 2079 TGCATCGGATATGAGCTGAGGCTTTGGTGGGAGAGAGAGTGTGCTGGGAAAGTTGA 2138
Dy 1187 TGGATCAGATCGGAGCTTAAGACTTAGAGGTGGAGGACGCGCTGTGTGGGACAGTTGA 1246
Qy 2139 GGTGAATCTCAGGCTGCGGTGGGAATTTCTGTGTCTGAATGGCTGGGGAATGAACATGTC 2198
Dy 1247 GGTGAGATTCAGAGACTGTTAGGAAGGTGTGTGACAGAGGCTGGGACCTGGAAGAAGC 1306
Qy 2199 TGAAGTTGTTGCAAGCAACTTGAATGTGGTCTGCAATAGGCTCTCCAGAGAGCCCTCA 2258
Dy 1307 TGATGTGTTTGCAGGCACTGGGATGGATCTGCACTCAAAACATCTTATCAAGTGA 1366
Qy 2259 TTTTACAGAAAGACATTACACATCTTAATGTGCAATTTCTGGCTGCACTGAGGGGAAGC 2318
Dy 1367 CTCCAAAATCCAGGCAACAAACATAGCTGTTTCTAAGTAGCTGTAAAGGAAATGAAC 1426
Qy 2319 CTCTCTCTGGGATGTATACAGTGGAGTGGAAACAGAGACTGCTGTCTCAATTAATATGGA 2378
Dy 1427 TTTCTTTGGGACTGCAAGAATGGCAATGGGTGGAGTACCTTGATCACTATGAAGA 1486
Qy 2379 AGCAAGTTTGATCTGCTCAGCCACAGGACCCAGGCTGGTGGAGCTGATATGCCCTG 2438
Dy 1487 AGCCAAAATTAACCTGCTCAGCCACAGGAAACCCAGAGCTGTTGGAGGGACATTCCTG 1546
Qy 2439 CTCTGGAGCTGTGAAGTGAACATGACAGACATAGGCGCTCTGTCTGTGATTTCTGATTT 2498
Dy 1547 TTTCTGGAGCTGTGAAGTGAAGCATGTGTGACACGTGGGGCTCCATCTGTGATTCGGAGCT 1606
Qy 2499 CTCTCTTCATCTGCCAATGTGTGTGAGAGAAATTAATTTGTGAGATGCTCATATCTCT 2558
Dy 1607 CTCTCTGGAAGCTGCCAGCGTTCTATGCGGGAATTAACAGTGTGGCAGAGTGTCTCTAT 1666
Qy 2559 TTCTGTGGAGATCACTTTGGAAGGAATGGTCTAACTTTGGGCGGAAAGTTCCAGTG 2618
Dy 1667 CTGGGGGAGCTCACTTTGGAGAGGAAATGGACAGATCTGGCTGAAGAAATTCAGTG 1726
Qy 2619 TGAAGGAGTGAACACTACCTTGATATGCCCATTTGTTCAACATCCGGAGACACTTG 2678
Dy 1727 TGAGGAGATGAGTCCCATCTTTCACTCTGCCAGTAGTAGCCCGCCGAGAGGAACPTG 1786
Qy 2679 TATCCACAGCAGAGAGTGTGAGTGTCTGTCCGATATACAGATGTCAGCTGTGCAAA 2738
Dy 1787 TAGCCACAGCAGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCCTGTTGGTGA 1846

Qy 2739 TGGCAAAATCCAGTGTGTGACGGGCAAGTGGAGATCAACGTGTGTGGACACTGGGGCTCACT 2798
Dy 1847 TGGCAAGACCCCGTGTGAGGCGAGAGTGGAGCTCAAAACGCTTGTGTGCTGGGATCCCT 1906
Qy 2799 GTGTGACACCCACTGGGACCCAGAGATGCCGTGTCTTATGACACAGCTCAGCTGTGG 2858
Dy 1907 CTGTAACCTCTCACTGGGACATAGAGATGCCATGTCTTTGCCAGCAGCTTAATAATGG 1966
Qy 2859 GACTGCTCTCTCAACACAGGAGGAAATATATTGGAAAGAACTGTTCTGTGTGGG 2918
Dy 1967 AGTTGCCCTTTCTACCCCGAGGAGGACAGCTTTTGGAAAGAAATGTCAGATCTGGAG 2026
Qy 2919 ACACAGTTTCATTGCTTTAGGAAATGAGTCACTTCTGTGATACTGCAAAATGACAGTTCT 2978
Dy 2027 GCATATGTTTCACTGCACTGGGACTGAGCAGACATGGAGATTTCTCTGTAACCTGCTCT 2086
Qy 2979 TGGAGCACTCCCTGTATCCATGGAATACTGTCTCTGTGATCTGCACAGGAGGCTGAC 3038
Dy 2087 AGGTGCTTTCATTATGTCTTCAGAGCAAGTGGCCTCTTAATCTGCTCAGGAAACAGTCT 2146
Qy 3039 CCAGCACTGTTTCCATGCTCGCAAAATGTATCTGACCCATATTTGCTGCAAGTTCCAGA 3098
Dy 2147 CCAACACTGCTCTGCTGCAATTCATGCTCTTTGGGCCCCAACAGGCTTACCACTCCAGA 2206
Qy 3099 GGGCAGTCTTTGATCTCTTTAGAGGACAAACGGCTCCGCTAGTGGATGGGACAGCGG 3158
Dy 2207 AGAAGTCTGTGGCTGATAGAGAGTGGTCACTTCCGCTGGTAAATGGAGGAGTGG 2266
Qy 3159 CTGTGCCGGGAGATGAGATCTATCAGACAGGGCTTCTGGGCCCCACCATCTGTGTGACGG 3218
Dy 2267 CTGTGCTGGGAGATGAGATCTATCATGAGGGCTCTTGGGCCCCACCATCTGTGTGACAG 2326
Qy 3219 CTGGCACTGAGCGATGCCACGCTGTGTCTCAAAAGCTGGCTGTGGAGTGGCTTCA 3278
Dy 2327 CTGGCACTGAGTGTATGCCACGCTGTGTGACAGAGCTGGCTGTGGAGAGGCTTAA 2386
Qy 3279 TGCCACAGTCTCTGCTCTACTTTGGGAGGGTCAAGGCCCATCTGGCTGGATGACCTGAA 3338
Dy 2387 TGCCACTGTTCTGCTCATTTTGGGAAAGAAACAGAGGCCCATCTGGCTGGATGAGATGA 2446
Qy 3339 CTGCACAGAGAGGAGTCCCACTTGTGGCAGTGGCTTCCCGGCTGGGGGACAGCA 3398
Dy 2447 ATGCAATGGAAGAAATCCCGCATTTTGGCAGTGGCTTCAACAGGCTGGGGGACGCA 2506
Qy 3399 CTGCAGGACAGAGGAGGACGAGGGTCACTGCTCAGAAATTCACAGGCTTTCAGGCTCTA 3458
Dy 2507 TTGCAGGACAGAGGAGGATCGGGAGTTATCTGCTCAGAAATTCATCTCTCAGACTGAC 2566
Qy 3459 CAGTGAACACTGAAACAGAGAGCTGTGTGGGAGATTTGAAAGTCTTCTATAGGGGACCTG 3518
Dy 2567 CAGTGAAGCCAGCAGAGAGGCTGTGACAGGGCTGTGGAAGTCTTGAAGTCTTACAAATGGA 2626
Qy 3519 GGGCAGCTCGGACAGGAGGAACATCACCACAGCCATAGCAGGCTTGTGTGACGACAGCT 3578
Dy 2627 GGGCACTGTTGGCAGAGATGATGCTGTGAAACCACTGTGGGTGGTGTGAGGACAGCT 2686
Qy 3579 GGGCTGTGGGAGATGAGTGTGTGAGCTGCGCCCTTTATCTAAGACAGGCTCTGTGTTT 3638
Dy 2687 GGGCTGTGCACAAAGAGGAAATCAACCTGTCATCTTTAGACAGGCAATCTCCATTC 2746
Qy 3639 CATGTGGTGGATGACATTCAGTGTCTTAAACGCAATATCTCCATATGGCAGTGGCTGTG 3698
Dy 2747 CATGTGGTGGCAATGTTTCACTGTGTCACAAAGGACCTGACAGCTGTGGCAGTGGCCATC 2806
Qy 3699 TGCCCCATGGGACCAAGAAATCTCCAGCCCCAGCAAGAGACCTGGATGATGATGGAAGA 3758
Dy 2807 ATCTCCATGGGAGAGAGACTGGCCAGCCCTCGGAGAGACCTGGATGATGATGGAAGA 2866
Qy 3759 TAGAATAAGAGTGTGTGGAGGACAGCCAGTGTCTCTGGGAGAGTGGAGATCTGGCAGCG 3818
Dy 2867 CAAGTAAGACTTTCAGGAAGGACCCACTTCTCTGTGAGAGTGTGGAGATCTGGCAGTGG 2926
Qy 3819 AGGCTCTCTGGGACAGTGTGTGATGACTCTCTGGGACCTGGCCGAGGCGGAAGTGGTGTG 3878

Db 647 CTTCAACATAGATGATGATCTGTCTATTTTGTAGACAACTTGAATGTGGAAGTGTGTCTGAC 706
Qy 1599 TGTGTTTGGTATGACCTATTTTAAAGAGCATCAGGACCTATTTGGCTGGATGAGTTTC 1658
Db 707 TTTCTCTGGTTCATCTAATTTTGGAGAGGCTCTGGACCAATCTGGTTTGTATGATCTTAT 766
Qy 1659 TTGCATTGGAAATGAGTCAATATCTGGGACTGTGAACACAGTGTGATGGGGAAGACATAA 1718
Db 767 ATGCAACGGAAATGAGTCACTCTGGAACATCAAGGATGGGGGAAGACATAA 826
Qy 1719 TTGTGTACACAGAGAGATGTATTTAACTGTCTCAGGTGATGCAACATGGGCGCTGAG 1778
Db 827 CTGTGTATGATCTGAGGATCTGAGGTGATTTGTCTAAAGGAGGACAGATCTGAGCCCTGAG 886
Qy 1779 GCTGGTGGGCGGACGAACCCCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTG 1838
Db 887 ACTGGTAGTGGAGTCACTGAATGTTTCAAGGAAGATTAGAAGTGAGATTCGAAGGAGATG 946
Qy 1839 GGGCACAGTGTGTGATGACGGCTGGAACAGTGAAGCTGACAGTGTGGGTGTGTAGCCAGCT 1898
Db 947 GGGGACAAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGGCATCAAGCAACT 1006
Qy 1899 GGACTGCCCATCTCTATCATTTGGCATGGGTCTGGGAACCTTCTACAGGATATGGAA 1958
Db 1007 GGGATGTCCAACTGCGCTCACAGCCATTTGGTCGAGTTAACCCCACTAAGGGATTTGGACA 1066
Qy 1959 AATTTGGCTCGATGATGTTTCTGTGATGGAGATGAGTCAGATCTCTGTCATGCAAGAA 2018
Db 1067 CATCTGGCTTACAGCGTTCTTCCAGGACATGAACCTGCTGTCTGGCAATGTAAACA 1126
Qy 2019 CAGTGGGTGGGGAATTAATGACTGCAGTACAGTGAAGATGTTGGAGTGTCTGTTCGA 2078
Db 1127 CCATGAATGGGGAAGCATTTATGCAATCACAATGAAGATGCTGGCGTGCATGTTCGA 1186
Qy 2079 TGCATCGGATGAGAGTGGGCTGTGGGTGGAAGCAGAGTGTGCTGGAAAGTTGA 2138
Db 1187 TGGATCAGATCTGGAGCTTAAGACTTAGAGTGGAGGACCGCTGTGCTGGGACAGTTGA 1246
Qy 2139 GGTGAATCTCCAGGTGCGGTGGGAATTTCTGTGTCTAATGGCTGGGGAATGAACATTGC 2198
Db 1247 GGTGGAGATTCAGAGACTGTTAGGAAGGTGTGTGACAGAGCTGGGACTGGAAGAAGC 1306
Qy 2199 TGAAGTTGTTTCAGGCAACTTGAATGTGGTCTGCAATCAGGCTCTCCAGAGAGCCCTCA 2258
Db 1307 TGATGTGTTTTCAGGCAAGCTGGGATGTGGATCTGCATCAAAACATCTTATCAAGTGA 1366
Qy 2259 TTTACAGAAAGACATACACATCTTAATCTCGAATTTCTGCTGCTAATGGCTGGGGAATGA 2318
Db 1367 CTCGAAATCCAGGCAACAAACATGCTGTTCTAAGTGTGTAACGGAAATGAAC 1426
Qy 2319 CTCTCTCTGGGATTTGATACGATGGGAGTGGAAACAGAGTCTGCTGCTCATTTAAATATGGA 2378
Db 1427 TTTCTTTTGGGACTGCAAGAACTGCAATGGGTGGGACTTACCCTGTGATCATATGAAGA 1486
Qy 2379 AGCAAGTTTGTCTCAGCCACAGGACGCCAGCTGGTGGAGCTGATATGCCCTG 2438
Db 1487 AGCCAAAATTAACCTCTCAGCCACAGGGAACCCAGACTGTTGGAGGGGACATTTCCCTG 1546
Qy 2439 CTCTGGAGCTGTGAAGTGAACATGACAGACATGGGCTGTCTGCTGTGATCTGATTT 2498
Db 1547 TTTCTGGAGCTGTGAAGTGAAGCTGGTGACACGTGGGGCTCCATCTGTGATTCGGACTT 1606
Qy 2499 CTCTCTTCATGCTGCCAATGTCTGTGAGAGAAATTAATTTGTGGAGATGCCATATCTCT 2558
Db 1607 CTCTCTGAGCTGCCAGCTTCTATGAGGGAATTAACGTGTGACAGTGTCTCTAT 1666
Qy 2559 TTCTGTGGGAGATCACTTTGGAAAAGGAATGGTCTAACTTTGGGCGGAAAAGTTCCAGTG 2618
Db 1667 CTTGGGGGAGCTCACTTTGGAGAGGGAATGGACAGATCTGGGCTGAAGATTTCCAGTG 1726
Qy 2619 TGAAGGAGTGAACACTCACTTGTGATATGCCCCATTTGTCACATTCGGGAAGACATG 2678
Db 1727 TGAGGAGATGATGCCCATCTTCTACTCTGCCAGTAGCACCCTCCGAGGAAGAACTTG 1786

Qy 2679 TATCCACAGCAGAGAAAGTTGGAGTTTGTCTGTCTCCGATATACAGATGTCCGACTTGTGAA 2738
Db 1787 TAGCCACAGCAGGAGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTGGTGA 1846
Qy 2739 TGGCAAAATCCAGTGTGACGGGCAAGTGAGATCAACGTGTGTGACACTGGGGCTCACT 2798
Db 1847 TGGCAAGACCCGCTGTGAGGCGAGAGTGGAGCTCAAAACGCTTGGTGTGGGATCCCT 1906
Qy 2799 GTGTGACACCCACTGGGACCCAGAGATGCCGCTGTCTATGTCACACAGCTCAGCTGCG 2858
Db 1907 CTGTAACCTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGAGCTTAATATGCG 1966
Qy 2859 GACTGCTCTCTCAACACAGGAGGAAAAATATATTGGAGAAAGAACTGTTCGTGTGGGG 2918
Db 1967 AGTTGCCCTTTCTACCCAGGAGGAGCAGTTTGGAAAAAGGAATGTGCAGATCTGGAG 2026
Qy 2919 ACACAGGTTTTCATTTAGGAAATGAGTCACTTCTGTGATCACTGTCAAAATGACAGTTCT 2978
Db 2027 GCATATGTTTCACTGCACTGGGACTGAGCAGCAGATTTGCTGTAACTGCTCT 2086
Qy 2979 TGGAGCACCTCCCTGTATCTCCATGGAATACTGTCTCTGTGATCTGCACAGGAAGCTGAC 3038
Db 2087 AGTGTCTTCAATATGCTCTCAGAGCAAGTGGCTCTGTATCTGCTCAGGAAACAGCTC 2146
Qy 3039 CCAGCCACTGTTTCCATGCCCTCGCAAAATGTATCTGACCCCATATTTGTCTGCAAGTTCCAGA 3098
Db 2147 CCAACACTGCTCTCGTCAATTCATGCTTTTGGGCCCAACAAAGGCTTACCATTCCAGA 2206
Qy 3099 GGGCAGCTCTTGTATCTCTTAGAGGACAAAGGCTCGGCTAGTGGATGGGAGCAGCGG 3158
Db 2207 AGAAGTCTGTGGGCTGCATAGAGAGTGGTCAACTTGGCTTGGTAAATGGAGAGGTCG 2266
Qy 3159 CTGTGCGGAGAGTAGAGATCTATCAGCAGCGCTTCTGGGGCACCATCTGTGATGACAG 3218
Db 2267 CTGTGCTGGGAGTAGAGATCTATCATGAGGCTCTTGGGGCACCATCTGTGATGACAG 2326
Qy 3219 CTGGACCTGAGCGATGCCACGCTGTGTGCTCAAAAGCTGGGCTGTGGAGTGGCTTCAA 3278
Db 2327 CTGGACCTGAGTATGCCACGCTGTTTGCAGACAGCTGGGCTGTGGAGAGGCCATAA 2386
Qy 3279 TGCACAGTCTCTGCTCACTTTGGGAGGGTGCAGGGCCCATCTGGCTGGATGACCTGAA 3338
Db 2387 TGCACACTGTTCTGCTCATTTTGGGGAAGGAACAGGGGCCCATCTGGCTGGATGAGATGAA 2446
Qy 3339 CTGCACAGAAAGGAGTCCCACTTGTGCGAGTGGCTTCCCGCGCTGGGGGACAGCA 3398
Db 2447 ATGCAATGGAAGAAGATCCCGCATTTGGCAGTGGCATTCACAGGCTGGGGGACAGCA 2506
Qy 3399 CTGCAGGCAACAGGAGGACGCGAGGCTCATCTGCTCAGAAATTCACAGCTTGGAGCTCTA 3458
Db 2507 TTGCAGGCAACAGGAGGATCGGGAGTTATCTGCTCAGAAATTCATGCTCTCAGACTGAC 2566
Qy 3459 CAGTCAAACTGGAACAGAGAGCTGTCTGGGAGATTGGAAGTCTTCTATAACGGGACCTG 3518
Db 2567 CAGTGAAGCCAGCAGAGAGGCTGTGCAAGGCTCTGGAAGTCTTCAATATGGAGCTG 2626
Qy 3519 GGGAGGCTGGGAGGAGGAACATCACACAGCCATAGCAGGATGTGTGCGAGGCT 3578
Db 2627 GGGACTGTTGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGCGAGGCT 2686
Qy 3579 GGGCTGTGGGAGAAATGGAGTTGTGAGCCTCGCCCTTTTATCTAAAGACAGGCTCTGTTT 3638
Db 2687 GGGCTGTGACAAAGGAAAAATCAACCTGCTCTTTTGAAGCAAGGCCATGTCCATTC 2746
Qy 3639 CATGTGGGTGATGATTCAGTGTCTTAAACGCAATATCTCCATATGGCATGTCTGTC 3698
Db 2747 CATGTGGGTGCAATTTTCACTGTCTCAAAAGGACTGACACGCTGTGGCAGTGGCCATC 2806
Qy 3699 TCCCCCATGGGAGCAAGAAATCTCCAGCCCAAGAGAGCTGATACATGTGACAA 3758
Db 2807 ATCTCCATGGGAGAGAGACTGGCCAGCCCCCTCGGAGGAGAGCTGGATACATGTGACAA 2866

Db 587 GTGTTCTGGAAGATAGAGATCAAAATCCAAGGACGGTGGGAACAGTGTGTGATCAATA 646
Qy 1539 ATGGAGCAACAAGAAATGAGCTGTGTGTGTAAACAATTTGGGATGTGGAAGCCCTATGCA 1598
Db 647 CTTCAACATAGATCATGATCTGTCTATTTGTAGACAACATTTGAATGTGNAAGTGTGTCTAG 706
Qy 1599 TGTGTTTGGTATGACCTATTTTAAAGAACATCAGGACCTATTTGGCTGGATGAGCTTTC 1658
Db 707 TTTCTCTGTTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGTTTGTATGATCTTAT 766
Qy 1659 TTGCATTTGGAATGAGTCAAAATATCTGGGACTGTGAACACAGTGTGATGGGGAACATAA 1718
Db 767 ATGCAACGGAATGAGTCAAGCTCTCTGGAATGTTCAGAAATCAAGATGGGGAAGCATAA 826
Qy 1719 TTGTGTACACAGAGAGGATGTGATTTAACTGCTCAAGTGTGATGAACATGGGGCTGAG 1778
Db 827 CTGTGATCATGCTGAGGATGCTGGAGTGTATTTGCTCAAGGAGGAGAGATCTGAGCCTGAG 886
Qy 1779 GCTGTGGGGCGGCAACACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTG 1838
Db 887 ACTGTAGATGGAGTCACTGAAATGTTCAGGAAGATTAGAAGTGAGATTCCAAGGAGAATG 946
Qy 1839 GGGCACAGTGTGTGATGAGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGAGCCAGCT 1898
Db 947 GGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGGCATGCAAGCAACT 1006
Qy 1899 GGACTGCCCATCTTCTATCAITGGCATGGCTCTGGAAACGCTTCTACAGGATATGGAAA 1958
Db 1007 GGGATGTCCAACCTGCGCTCAGACCATTTGGTCGAGTTAACGCCAGTAAGGGATTTGGACA 1066
Qy 1959 AATTTGGCTCGATGATGTTTCTGTGTGATGGAGATGAGTCAGATCTCTGCTGTCAGGAA 2018
Db 1067 CATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGCTGTGCAATGTAAACA 1126
Qy 2019 CAGTGGGTGGGAAATATGACTGCAGTCACAGTCAAGATGTTGAGTGATCTGTCTCA 2078
Db 1127 CCATGAATGGGGAAGCATTTATGCAATCACAATGAAGTGTGGCGTGACATGTTCTGA 1186
Qy 2079 TGCACTCGATATGGAGCTGAGGCTTTGTGGTGGAAAGCAGCAGCTGTGTGGAAGTTGA 2138
Db 1187 TGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGACGCCCTGCTGTGGACAGTTGA 1246
Qy 2139 GGTGAATGTCAGGTCGCGTGGGAATTTCTGTGTCTAATGGCTGGGAATGAACATGTC 2198
Db 1247 GGTGAGATTCAGACACTGTTAGGGAAGCTGTGCACAGAGCTGGGACTCAAAAGAC 1306
Qy 2199 TCAAGTCTTTCAGGCAACTGTATGTGGTCTGCAATCAGGCTCTCCAGAGACCTCA 2258
Db 1307 TGATGTGTTTTCAGGCACTGGGATGTGGATCTGCACCTCAAAACATCTTATCAAGTGTGA 1366
Qy 2259 TTTTCAGAAAGAACATTAACATCTTAATGTCGAATTTCTGGCTGCACCTGGAGGGGAAGC 2318
Db 1367 CTCGAAATCCAGGCAACAACATAGGCTGTTCTAAGTAGCTGTAAAGGAAATGAAC 1426
Qy 2319 CTCCTCTGGGATGTATACGATGGAGTGGAAACAGACTGCGTGTCTAATTTAAATATGA 2378
Db 1427 TTTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGTATCACTATGA 1486
Qy 2379 AGCAAGTTTGTCTCTCAGCCCAAGGAGCCAGGCTGGTGGAGCTGATATGCCCTG 2438
Db 1487 AGCCAAAATTTACCTGCTCAGCCCAAGGGAACCCAGACTGGTGGAGGGACATTTCCCTG 1546
Qy 2439 CTCGAGCTGTTGAAGTGAACATGCACACATGGGCTCTGCTGTGATTTGATTT 2498
Db 1547 TTTCTGGAGCTGTTGAAGTGAACATGGTGACACGTTGGGGCTCCATCTGTGATTCGGACTT 1606
Qy 2499 CTCCTCTCATGTCGAATGTCTGTGAGAGAAATTAATTTGGAGATGCCATATCTCT 2558
Db 1607 CTCCTGGAAGCTGCCAGGTTCTATGCAGGAAATTAACAGTGTGCACAGTTGTCTCTAT 1666
Qy 2559 TTTCTGGGAGATCACTTTGGAAGGAATGGTCTAATCTGGGCGCAAGATTTCCAGTG 2618
Db 1667 CCTGGGGGAGCTCACTTTGGAGAGGAAATGGACAGATCTGGGCTGAAGAATTTCCAGTG 1726

Qy 2619 TGAAGGAGTGAAGTCACTTTGCAATATATGCCCATTTGTTCAACATCCCGAAGACACTTG 2678
Db 1727 TGAGGACATGAGTCCCATCTTTTCACTCTGCCAGTAGCACCCGCCAGGAAGAACTTG 1786
Qy 2679 TATCCACAGCAGAGAGTTGGAGTTGTCTGTGTTCCCGATATACAGATGTCGGACTTTGAA 2738
Db 1787 TAGCCACAGCAGGAGTGTGGAGTAGTCTGTCTCAAGATACACAGAAATTCGCTTGTGAA 1846
Qy 2739 TGGCAATCCCACTGTGAGGGCAAGTGGAGATCAAGTGTGAGACACTGGGCTCACT 2798
Db 1847 TGGCAAGACCCGCTGTGAGGGCAGAGTGGAGCTCAAAACGCTTGTGCTGGGATCCCT 1906
Qy 2799 TGTGACACCCACTGGGACCCAGAGATGCCGCTGTCTATGCAGACAGCTCAGCTGTG 2858
Db 1907 CTGTAACTCTACTGGACATAGAAGATGCCCATGTCTTTGCCACGACTTAAATGTG 1966
Qy 2859 GACTGCTCTCAACACAGGAGGAAATATATTTGGAGAAAGATGTTGCTGTGTGGGG 2918
Db 1967 AGTTGCCCTTTTACCCAGGAGGACAGCTTTTGGAAAAGGAAATGGTCAGATCTGGAG 2026
Qy 2919 ACACAGGTTTCACTGCTTAGGGAATCAGTCACTTCTGGATAACTGTCAATGACACTCT 2978
Db 2027 GCATATGTTTCACTGCTAGGACTGAGCAGCAGTGGAGATTTGCTGTAACTGCTCT 2086
Qy 2979 TGGAGCACCTCCCTGTATCCATGGAAATACTGTCTGTGATCTGCACAGGAAGCCTGAC 3038
Db 2087 AGGTGCTTCAATATGTCTTTCAGAGCAAGTGGCTCTGTAACTGCTCAGGAACCACTC 2146
Qy 3039 CCAGCCACTGTTTCCATGCTCGCAATATATGTACCCATATTTGCTGTGAGTTCCAGA 3098
Db 2147 CCAACACTGCTGCTGTCATTCATCTGTTTGGCCCAACAAAGCCCTACCATTCCAGA 2206
Qy 3099 GGGCAGTGTGTGATCTGCTTAGAGGACAAACGGCTAGTCCGCTAGTGGAGGACAGCG 3158
Db 2207 AGAAGTGTGTGGCTGCTAGAGAGTGTCAACTTCCGCTGTGAATGGAGAGGTG 2266
Qy 3159 TTTGTCGGGAGAGTAGAGATCTATCACAGCAGCTTCTGGGACCATCTGTGATGACGG 3218
Db 2267 CTGTGCTGGGAGTAGAGATCTATCATGAGGCTCTCTGGGACCATCTGTGATGACAG 2326
Qy 3219 CTGGGACCTGAGGATGCCACAGTGTGTCTCAAGAGTGGGCTGTGGAGTGGCTTCAA 3278
Db 2327 TTTGGGACCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2386
Qy 3279 TGGCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3338
Db 2387 TGGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2446
Qy 3339 TTTGCAAGCAAGGAGTCCACTTTGTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3398
Db 2447 ATGCAATGGAAGAAATCCCGCATTTGGCAGTGGCCATTCACAGGCTGGGGCAGCAAA 2506
Qy 3399 CTGAGGCAACAAGGAGGACGCGAGGTCTATCTGCTCAGAAATTCACAGCCTTGAGGCTCTA 3458
Db 2507 TTTGAGGCAACAAGGAGGATGCGGAGTTATCTGCTCAGAAATTCATGCTCTGAGACTGAC 2566
Qy 3459 CAGTGAACCTGAACAGAGAGCTGCTGCTGGAGATTTGGAAGTCTTCTATACGGGACCTG 3518
Db 2567 CAGTGAACCTGAACAGAGAGCTGCTGCTGGAGGCTCTGGAAGTCTTCTATACAGGCTGCT 2626
Qy 3519 GGGCAGGCTGCGCAGGAGGAACATCACACAGCAGGCAATGTGTGAGGAGGAGCT 3578
Db 2627 GGGCAGTGTGGCAGAGATGAGTGTCTGAAACACTGTGGGTGTGGTGTGGAGGAGCT 2686
Qy 3579 GGGCTGTGGGAGATTTGAGTGTCTGAGCCTCGCCCTTTATCTAAGACAGGCTCTGGTTT 3638
Db 2687 GGGCTGTGGCAGACAAGGGAATCAACCTGCACTTTTAGACAAGGCCATGTCCATTC 2746
Qy 3639 CATGTGGGTGGATGACATTCAGTGTCTTAAAGGCAATATCTCCATATGCGAGTGCCTGTC 3698
Db 2747 CATGTGGGTGGCAATGTTTCAAGTGTCCAAAGGAGGCTGACAGGCTGTGGCAGTGCCCATC 2806

[illegible]

1658	CTTGCAATTGGAAATCAGTCAAAATATCTGGGACTCTGAACACAGCTGGATGGGGAAGCATA	1717
767	TATGCAACGGAAATGAGTCAGCTCTCTGGAACATCAAAACATCAAGGATGGGGAAGCATA	826
1718	ATTGTGTACACAGAGAGATGTGATTGAACCTGCTCAGGTGATGCAACATGGGCGCTGA	1777
827	ACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAGATCTCAGCGCTGA	886
1778	GGCTGGTGGCGGAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGT	1837
887	GACTGGTAGATGGAGTCACTCAATGTTTCAGGAAGATTAGAAGTGAGATTTCCAAGGAGAAT	946
1838	GGGCGACAGTGTGTGATGACGGCTGGAAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGC	1897
947	GGGGACAATATGTGATCACGGCTGGGACAGTTACGATGCTGCTGTGGCATGCAAGCAAC	1006
1898	TGGACTGCCCATCTTCTATCATTTGGCATGGGTCTGGGAACGGCTTCTACAGGATATGGAA	1957
1007	TGGGATGTCCAACCTGCCGTACAGCCCAATGGTTCGAGTTAAACGCCAGTAAGGGATTTGGAC	1066
1958	AAATTGGCTCGATGATGTTTCTGTGATGGAGATGAGTCAGATCTCTGTGTCATGCAGGA	2017
1067	ACATCTGGCTTGCACAGCGTTTCTTGGCCAGGGACATGAACCTGCTGTCTGGCAATGTAAAC	1126
2018	ACAGTGGGTGGGGAATAATCACTGCAGTCAACATGAAGATGTTGGAGTGATCTGTCTCTG	2077
1127	ACCATGAATGGGGAAGCAATTATTTGGCAATCACATGAAGATGCTGGCGTGCATGTTCTG	1186

QY 2078 ATGCATCGGATATGGAGCTTGTGGGTGGGAAGCAGCAGGCTGTCTGGAAAGATTG 2137
Db 1187 ATGGATCAGATCTGGAGCTTAAGACTTAGAGGTGGAGCGCCCTGTCTGGACAGATTG 1246
QY 2138 AGGTGAATCTCAGAGTGCCTGGGAATCTCTGTCTTAATGGCTGGGAAATGAACATTG 2197
Db 1247 AGGTGAGATTCAGAGACTGTAGGAAGGTGTGTACAGAGCTGGGACTCAAGAAAG 1306
QY 2198 CTGAAGTTCTTTCAGGCAACTTGAATGTGGGTCTCAATCAGAGTCTCCAGAGCCCTC 2257
Db 1307 CTGATGTGTTTCAGGACAGCTGGGATGTGGATCTGCATCAAAACATCTTATCAAGTGT 1366
QY 2258 ATTTCAGAAAGAACATTACACATTTAAATGTGCAATTTCTGGCTGCATCGGAGGGAAG 2317
Db 1367 ACTCCAAAATCAGGCAACAAACACATGCTGTTCTAAGTAGCTGAACGGAATGAAA 1426
QY 2318 CTTCTCTCTGGGATCTATACGATGGAGTGGAAACAGACTGCGTCTCATTTAAATATGG 2377
Db 1427 CTTCTCTTTGGGACTTCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTATGAAG 1486
QY 2378 AAGCAAGTTTGTATCTGCTCAG--CCACAGGCGAGCCAGGCTGGTGGAGCTGATATGCC 2435
Db 1487 AAGCCAAAATTAACCTGCTCAGGCCACAGGGNACCCACACTGGTGGAGGGGACATTC 1546
QY 2436 CTGCTCTGACGCTGTGAAGTGAACATCGACACATCGCCGCTCTGTCTGTGATTCTGA 2495
Db 1547 CTGTTCTGACGCTGTGAAGTGAACATGGTGACACGCTGGGGCTCCATCTGTGATTCTGA 1606
QY 2496 TTTCTCTCTTCATCTGCGCAATGTGCTGTGCAGAGAAATTAATTTGGAGATGCCATATC 2555
Db 1607 CTTCTCTCTGGAAGCTGCCAGGCTTCTATGACAGGAATTAACAGTGTGGCACAGTTGCTC 1666
QY 2556 TCTTTCTGTGGGAGACTCTTTGGAAAGGGAATGCTTAACCTTGGGCCGGAAGATTCCA 2615
Db 1667 TATCTTGGGGAGCTCACTTTGGAGAGGGAATGGACAGACTTGGGCTGMAAATTTCA 1726
QY 2616 GTGTGAAGGAGTGAACACTCACTTTGCAATATGCCCATTTGTTCAACATCCGGAAGACAC 2675
Db 1727 GTGTGAGGACATGAGTCCCATCTTCACTCTGCCAGTAGCACCCGCCAGAGGAAC 1786
QY 2676 TTGTATCCACAGCAGAGAAATGGAGTTGCTGTGTTCCCATATACAGATGTCGGACTTGT 2735
Db 1787 TTGTAGCCACAGCAGGGAATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTGT 1846
QY 2736 GAATGCAAAATCCAGTGTGAGGCGCAAGTGGAGATCAACGCTCTTGGACACTGGGGCTC 2795
Db 1847 GAATGCAAGACCCCTGTGAGGCGAGAGTGAGCTCAAAACGCTTGGTGCCTGGGGATC 1906
QY 2796 ACTGTGTACACCCACTGGGACCCAGAGATGCCGCTGTCTATGACAGACGCTCAGCTG 2855
Db 1907 CCTCTGTAACCTCTCACTGGGACATAGAAGATGCCATGTTCTTTGCCAGCAGCTTAAATG 1966
QY 2856 TGGGACTGCTCTCAACACAGGAGAAATATATTGGAGAAGAAGTTGCTGTGTG 2915
Db 1967 TGGAGTTGGCCCTTTACCCCGAGGAGGACGCTTTTGGAAAGGAAATGGTCAGATCTG 2026
QY 2916 GGACACAGGTTTCATTTAGGGAATGAGTCACTTCTGGATAACTGTCAAAATGACAGT 2975
Db 2027 GAGGCATATGTTCACTGCACTGGGACTGAGCAGACATGGGAGATTGCTCTAATCTC 2086
QY 2976 TCTTGGAGCACTTCCCTGTATCCATGGAATATCTGTCTCTGTGATCTGCACAGGAAGCCT 3035
Db 2087 TCTAGTGTCTTATATGTTCTTACAGCAAGTGGCTCTGTAACTCTGCTCAGGAAACCA 2146
QY 3036 GACCCAGCCACTGTTTCCATGCTCGCAATGTATCTGACCCNATATTGTCTGCACTTCC 3095
Db 2147 GTCCCAAAACATGCTCTCTGTGCAATTCATCTGCTTTTGGGCCCAACAGGCCCTACCATTCC 2206
QY 3096 AGAGGCCACTGCTTTCATCTGCTTAGAGCAAAACCGCTCCGCTAGTGGATGGGACAG 3155
Db 2207 AGAAGAAAGTGTCTGGCTGCTGATAGAGAGTGTCAACTTCCCTGTTAAATGGAGAGG 2266
QY 3156 CCCTGTGCGGGAGAGTAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGA 3215

Db 2267 TCCCTGTCTGGGAGAGTAGATCTATCATGAGGCTCTCTGGGGCACCATCTGTGATGA 2326
QY 3216 CGCTGGGACCTTGAGGATGCCACGTTGTGTCTCAAAAGCTGGCTGTGGAGTGCCTT 3275
Db 2327 CAGCTGGGACCTTGAGTATGCCACGTTGTGTGACAGACAGCTGGGCTGTGGAGAGCCAT 2386
QY 3276 CAATGCCACGCTCTCTGCTCACTTTTGGGAGGGGTCCAGGCCCATCTGCTGTGATGACCT 3335
Db 2387 TAATGCCACTGTTCTGCTCATTTTGGGAAGGAACAGGCCCATCTGCTGTGATGAGAT 2446
QY 3336 GAATGTCACAGAACCGAGTCCACATTGTGGCAGTGCCTTCCCGGGCTGTGGGGGACGA 3395
Db 2447 GAAATGCAATGGAAGAAATCCCGCATTTGGCAGTGCCTTACACGGCTGGGGGACGA 2506
QY 3396 CGACTCGCGCAACAGGAGGAGCGAGGTCTCTCTCAGAAATTCACAGCCTTTGAGGCT 3455
Db 2507 AATTCGACGCAAGAGGATGCGGGAGTTATCTCTCAGAAATTCATGCTCTGAGACT 2566
QY 3456 CTACAGTGAACCTGAAACACAGAGAGCTGTCTGGGAGATTGGAAGTCTTCTATAACGGAC 3515
Db 2567 GACCACTGAAGCCAGCAGAGAGGCTGTGCAGGCGCTCTGGAAGTTTTCACATGAGC 2626
QY 3516 CTGGGCAAGCTGCGGAGGAGGAACATCACACAGCCATAGCAGGATTTGTGTGACAGCA 3575
Db 2627 TTGGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGTGTGACAGCA 2686
QY 3576 SCTGGGCTGTGGGAGAATGGAGTTGTGACGCTCGCCCTTTATCTTAAGACAGGCTCTGG 3635
Db 2687 GTTGGGCTGTGCACAGAAAGGGAATCAACCTGCTATCTTAGAAGAGCCCATGTCAT 2746
QY 3636 TTTTATGTGGGTGATGACATTCAGTGTCTTAAACGCATATCTCCATAT-GGCAGTGCC 3694
Db 2747 TCCCATGTGGGTGACAAATGTTTCAAGTGTCCAAAGGACCTGACACGCTGTGGCAGTGCC 2806
QY 3695 TGTCTCCCATGGGAGCGAAGAAATCTCAGCCAGCAGAGAGAGCTTGGATCAGATGTG 3754
Db 2807 CATCATCTCCATGGGAGAAGAGACTGGCCAGCCCTCGGAGGAGAGCTTGGATCAGATGTG 2866
QY 3755 AGATAGAAATAGAGTGGTGGAGGAGACACCGAGTCTCTGGAGAGTGGAGATCTGCG 3814
Db 2867 AACAAAGATAAGACTTCAGGAAGGACCCATCTTCTGTCGAGCTGTGGAGATCTGCG 2926
QY 3815 ACAGAGCTCTTGGGCGACAGTGTGTGATGACTCTCTGGGACCTGGCCGAGCGGAAGTGG 3874
Db 2927 ATGGAGTTCTTGGGGACAGTGTGTGATGACTTTTGGGACTTTGGAGATGCTCAGGTGG 2986
QY 3875 TGTGTGACAGCTGGGCTGTGGCTGTCTGCTGCTGCTGCTGAGGAGAGCTTCTGTTTGGCC 3934
Db 2987 TGTGTCAACAACCTTGGCTGTGTGCTTGAAGCATTTCAAGAGCAGAGATTGTTGGTC 3046
QY 3935 AGGAACTGGAACCATCTGTTGGATGACATGGGTGCAAGGAAATGAGTCAATTTCTAT 3994
Db 3047 AGGGACTGAGCCGATATGGCTCAATGAAGTGAAGTGAAGGAATGAGTCTTCCTTGT 3106
QY 3995 GGGACTGTCCGCAAAACCTTGGGACAGAGTACTGTGACACAAAGGAAGTGTCTGGCG 4054
Db 3107 GGGATTGCTTCCGACAGCTGGGGCCATAGTGAGTGTGGGCAACAAGGAAGCTGACAG 3166
QY 4055 TGAGTGTCTTGGACAGTCTGCTGAA-----ATCACTGAATGCTCTCT 4096
Db 3167 TGAATTTGCACAGATATTTTCAAGTGCAGAAACCCCAAAAGCCACAACAGGCTGCTCAT 3226
QY 4097 CAGGTCAATTAGCATTATTTATCCAGTATCTTTGGGCTCTCTCTCTCTGTTCTCTTTA 4156
Db 3227 CCGGTCAAGTCACTCTTTATTTGAGTGGGATCTCTGGGTGTTCTGTTGGCCATTTTCG 3286
QY 4157 TTTCTATTTCTCAGCTGGTCCGAGTTTCAGAAACAAAACATCTGCCCTCTCAGAGTTTCAA 4216
Db 3287 TCGCATATTTCTTCTTGAATTAAGGAGAGAGACACAGAGACACAGAGAGCTTCTCTCT 3343
QY 4217 CCAGAGGAGGGGTCTCTTCGAGGAGAAATTTATTCATGAGATGGAGAGCTGCTCTCAAG 4276

QY 2019 CAGTGGGTGGGAAATTAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATCTGTTCTCA 2078
DB 1127 CCATGAATGGGGAAGCAATATTGCAATCACAATGAGATGCTGGCGTGACATGTTCTGA 1186
QY 2079 TGCATCGGATATGGAGCTGAGCGTTGTGGTGGAGAGCAGAGGTGTGCTGGGAAAGTTGA 2138
DB 1187 TGGATCAGATCTGGAGCTAAGACTTAGAGTGGAGCAGCCCTGTGCTGGGACAGTTGA 1246
QY 2139 GGTGAATGCCAGGTGCCGTGGGAATTTCTGTGCTAATGCTGGGGAATCAACATGTC 2198
DB 1247 GGTGAGATTCAGAGACTGTTAGGGAAGGTGTGACAGAGGCTGGGACTGAAAGAAGC 1306
QY 2199 TGAAGTGTGTCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTGCAGAGAGCCTCA 2258
DB 1307 TGAATGTGTCAGGCAAGTGGATGAGATGAGATCTGCATCAAAACATCTTATCAAGTGA 1366
QY 2259 TTTACAGAAAGAACATTACACATCTTAATGTCGAATTTGCTGCATCGAGGGGAAGC 2318
DB 1367 CTCCAAAATCCAGGCAACAAACACATGCTGTTCTAAGTAGCTTAACGGAAATGAAC 1426
QY 2319 CTCTCTCTGGGATTTATACGATGGAGTGGGAACAGACTGCTGCTCATTTAAATATGA 2378
DB 1427 TTCTCTTTGGGACTGCAGAACTGGCAATGGGGTGGACTTACTGTGATCACTATGAAGA 1486
QY 2379 AGCAAGTTTGATCTGCTCAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2438
DB 1487 AGCCAAAATACCTGCTCAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1546
QY 2439 CTCTGAGGTGTTGAAGTGAACATGCAGACACATGAGGAGGAGGAGGAGGAGGAGGAGG 2498
DB 1547 TTTCTGAGGTGTTGAAGTGAACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1606
QY 2499 CTCTCTGATGCTGCAATGCTGTCGAGAGAAATTAATTTGAGAGATGCCATATCTCT 2558
DB 1607 CTCTCTGAGGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1666
QY 2559 TTTCTGAGGAGTACTTTGAAAAGGGAATGGTCTAATCTGGGCGGAAAGGTTCCAGTG 2618
DB 1667 CCTGGGGAGGAGTACTTTGAGAGGGAATGGAGAGATCTGGGCTGAAGATTTCCAGTG 1726
QY 2619 TGAAGGAGTGAACATCACTTTCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2678
DB 1727 TGAGGAGATAGTCCCATCTTCTACTCTGCGCAGTAGCACCCCGCCAGAGAACTTG 1786
QY 2679 TATCACAGCAGAGAGTGGAGTTGCTG ----- 2708
DB 1787 TAGCCACAGCAGGAGTGTGGAGTAGTCTGCTCAAGTAGACCCAGAAACATCTTTAAT 1846
QY 2709 ----- 2708
DB 1847 TGGTTCTTATCTGTGAAGGGACAGGTTAGGGAGTCATAGCTGTCTTTTCTAAAGCC 1906
QY 2709 -----TTCCCATATACAGATGCTGGAGTGTGAATGGCAATCCAGTGTGACGG 2759
DB 1907 CTGCTCTCTCCAGATACACAGAAATTCGCTTGGTGAATGGCAAGACCCCGTGTGAGG 1966
QY 2760 GCAAGTGGAGATCAAGCTGTTGGACACTGGGGCTCAGCTGTGACACCCACTGGGAGCC 2819
DB 1967 CAGAGTGGAGTCAAAACGCTTGGTGGCTGGGGATCCCTCTCTAATCTCACTGGGACAT 2026
QY 2820 AGAAGATCCCGTGTCTATGCAGACAGCTCAGCTGTGGAGTGTCTCTCAACACAGG 2879
DB 2027 AGAAGATCCCGATGTTCTTGGCAGCAGCTTAAATGTGGAGTTGCCCTTTCTACCCAGG 2086
QY 2880 AGGAAATATATGGAGAAAGAGTGTGCTGTGGGACACAGGTTTCATGCTTAGG 2939
DB 2087 AGGACACGTTTGGAAAAGAAATGGTGCAGATCTGGAGGCATATGTTTCACTGCACTGG 2146
QY 2940 GAATGAGTCACTTCTGGATACCTGTCMAATGACAGTCTTTGGAGCACCTCCCTGTATCA 2999
DB 2147 GACTGAGCAGCACATGGGAGATTGCTCTAATCTGCTAGGTGCTTCAATATGCTCTTC 2206

QY 3000 TGAATAATACTGTCTGTGATCTGCACAGGAAGCCTGACCCAGCAGCTGTTCTTCATGCT 3059
DB 2207 AGAGCAAGTGGCCTCTGTAATCTGCTCAGGAARACCAGTCCCAACACACTGCTCTCGTGA 2266
QY 3060 CGCAATATGATCTGACCCATATTTGCTGACATTTCCAGAGGCGAGTGTGTTGATCTGCT 3119
DB 2267 TTTATCTGCTTTTGGGCCCAACAAAGGCTACCATTTCCAGAAAGAAAGTGTGTTGCTGCA 2326
QY 3120 AGAGGACAAAGGCTCCGCTAGTAGTGGGACAGCCGCTGTGCGGGAGAGTAGAGAT 3179
DB 2327 AGAGAGTGGTCAACTTCGCTTGAATGGAGAGGTGCTGTGCTGGGAGTAGTAGAT 2386
QY 3180 CTATCAGCAGGCTTCTGGGCGCACCATCTGTGATGACGGCTGGGACCTGAGGCATGCCCA 3239
DB 2387 CTATCATGAGGCTCTGGGCGCACCATCTGTGATGACAGCTGGGACCTGAGTGTGATGCCA 2446
QY 3240 CGTGTGTGTCAAAAGCTGGGCTGTGGAGTGGCTTTCAATGCCACGGTCTCTGCTCACT 3299
DB 2447 CGTGTGTGTGCACAGCTGGGCTGTGGAGAGGCCATTAATGCCACTGTTCTGCTCATTT 2506
QY 3300 TGGGAGGGGTCCAGGCGCCATCTGGCTGGATGACCTGGAATGCACAGGAACGAGTCCCA 3359
DB 2507 TGGGAAGGAACAGGGGCCCATCTGGCTGGATGAGATGAATGCAATGGAAAGAAATCCCG 2566
QY 3360 TTTGTGGCAGTCCCTTCCCGGCTGGGGCGAGCAGCAGTCTCAGGSCACAAGGAGGAGC 3419
DB 2567 CATTTGGCAGTCCCATTCACACGGCTGGGGCGAGCAAAATTTGAGSCACAAGGAGATGC 2626
QY 3420 AGGGTGTCTCTCAGAAATTCAGACCTTTGAGGCTCTACAGTGAATGAACAGAGAG 3479
DB 2627 GGGAGTTATCTCTCAGAAATTCATCTCTGAGACTGACAGTGAAGCCAGCAGAGAGGC 2686
QY 3480 CTGTCTGGAGATTTGGAAGTCTTCTATAACGGGACCTGGGCGAGCTCGGCGAGGAGAA 3539
DB 2687 CTGTGAGGCGCTCTGGAAGTTTTTACAAATGAGCTTGGGCGACTGTTGGCGCAAGTAG 2746
QY 3540 CATCACACAGCCTATGAGGAGCTTGTGTCAGGAGCTGGGCTGTGGGAGAAATGGAGT 3599
DB 2747 CATGTCTGAAACACTGTGGGTGTGTGTCAGGAGCTGGGCTGTGTCAGACAAAGGAA 2806
QY 3600 TGTAGCCTCGCCCTTTATCTAAGACAGGCTCTGTTTCATGTGGTGGATGACATCA 3659
DB 2807 AATCAACCTTGATCTTTAGACAAGGCCATGTCATTTCCATGTGGGTGGCAATGTTCA 2866
QY 3660 GTGCTCTAAAGCAGATATCTCCATATGGCAGTGCCTGTGCCCCATGGAGCGGAGAAAT 3719
DB 2867 GTGCTCAAAAGACCTTGACACGCTGTGGCAGTGGCCATCATCTCCATGGGAAAGACT 2926
QY 3720 TTTCCAGCCCCAGAGAGACCTGGATCATATGTGAAGATAGAATAAGAGTCCGTGGAGG 3779
DB 2927 GGGCAGCCCTCGGAGGAGACCTGGATCATATGTGACAAAGATAAGACTTTCAGGAAGG 2986
QY 3780 AGACACCGAGTCTCTGGGAGAGTGGAGATCTGGCAGCAGGCTCTCTGGGGCACAGTGTG 3839
DB 2987 ACCCACTTCTCTTCTGGAGCTGTGGAGATCTGGATCTGGATCTGGGCGACAGTGTG 3046
QY 3840 TGTAGCTCTGGGACCTGGGCGAGGCGAAGTGTGTCTCAGCAGCTGGGCTGTGGCTC 3899
DB 3047 TGTAGCTCTGGGACTTGGACGAGTCTCAGGTGTGTCTCAACACTTGGGCTGTGGTCC 3106
QY 3900 TGTCTGCTGCTGAGGAGCCTGTGCTGTTGGCAGGGAATGGAACATCTGTTGGA 3959
DB 3107 AGCTTTGAAAGCATTTCAAGAAAGCAGAGTTTGTGTCAGGGGACTGGACCATATGGCTCAA 3166
QY 3960 TGCATGCGGTCCAAAGAAATGAGTCAATTTCTATGGGACTCTCAGCCCAACCTGGGG 4019
DB 3167 TGAAGTGAAGTCCAAAGGAAATGAGTCTTCTTGTGGGATTTCTCTGGCAGAGCTGGGG 3226
QY 4020 ACAGAGTGAAGTGTGGACACAAAGGAGATCTGGCTGGGTGAGTGTCTGGACAGTCTGAA 4079
DB 3227 CCAATAGTGTGGGACAAAGAGAGCTCCAGTGAATGCAAGATATTTCAAGTGA 3286
QY 4080 -----ATCACTGAATGCTCTCCTCAGGTCAATTTAGCACTTTATTTATC 4121


```
QY 3070 TCTGACCCATATTTCTCTCAGTTTCCAGAGGCGACTGCTTTGATCTGCTTAGAGACAAA 3129
      ||||| |
Db 595 TCTGAAGAGCAGGCTCTGTAGTCTCAGAAGAGATCGCCCCCTACTGCTCAGATAGCAGA 654
      ||||| |
QY 3130 CGGCTCCGCCCTAGTGGAGTGGGACAGCCGCTGTGCGGGAGAGTAGAGATCTATCACGAC 3189
      ||||| |
Db 655 CAGCTCCGCCCTGGTGGAGCGGAGAGTCCCTGTGCGGGAGAGTAGAGATCTATCACGAC 714
      ||||| |
QY 3190 GGTCTCTGGGGCACCATTCTGTGATCAGCGCTGGGACCTGAGCGATGCCACGTGCTGTGT 3249
      ||||| |
Db 715 GGTCTCTGGGGCACCATTCTGTGATCAGCGCTGGGACCTGAGCGATGCCACGTGCTGTGT 774
      ||||| |
QY 3250 CAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCAGCGTCTCTGCTCACTTTTGGGAGGG 3309
      ||||| |
Db 775 AGGCAGTTGGGCTGTGGAGAGCCCTCAATGCCAGCGTCTCTGCTCACTTTCGGGCGCAG 834
      ||||| |
QY 3310 TCAGGGCCCATCTGCTGATGATGACCTGAACCTGACAGAGAGGATGCCACTTGTGGGAG 3369
      ||||| |
Db 835 TCAGGGCCCATCTGCTGATGATGACCTGAACCTGACAGAGAGGATGCCACTTGTGGGAG 894
      ||||| |
QY 3370 TGCCCTTCCCGGGCTGGGGCAGCAGCAGCTGCGAGCAGCAGGAGGAGCGAGGGCTCATC 3429
      ||||| |
Db 895 TGCCCTTCCCGGGCTGGGGCAGCAGCAGCTGCGAGCAGCAGGAGGAGCGGGGCTCATC 954
      ||||| |
QY 3430 TGCTCAGAAATTCACAGCCCTTGAGGCTCTACAGTGAACCTGAAACAGAGAGCTGTCTGG 3489
      ||||| |
Db 955 TGCTCAGAAATTCACAGCCCTTGAGGCTCTACAGTGAACCTGAAACAGAGAGCTGTCTGG 1008
      ||||| |
QY 3490 AGATTGGAGTCTTCTATAACGGGACCTGGGCGACCGCTGGGAGGAGGAATCACACACA 3549
      ||||| |
Db 1009 TGGCTGGAAGTTTCTACAAATGGGACCTGGGCGACCGTGTCCGAAAGCCCATGGAAG 1068
      ||||| |
QY 3550 GCATAGCAGGCAATGTGTGAGGAGCAGTGGGCTGTGGGAGAAATGGAGTCTCAGCCTC 3609
      ||||| |
Db 1089 ATCACTGTGCTCATCTGCACTGCTGCTGCTGGGAGAGTGAAGTCTCAACTCT 1128
      ||||| |
QY 3610 GCCCTTTATTAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTCAGTGTCTCTAAA 3669
      ||||| |
Db 1129 TCTGTGTCTTAGAAGAGGTTTATAGACTCGTGGTGGTAGATGGAATCCGGTGTGCGA 1188
      ||||| |
QY 3670 AGCATATCTCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3729
      ||||| |
Db 1189 ACAGATATCTCTCTGCGAGTGTCCATCTGACCATTTGAATATATCTATCTCTCTCA 1248
      ||||| |
QY 3730 GCAGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3783
      ||||| |
Db 1249 AAGAGAGAGCTTATCTATCTGATGATGATGATGATGATGATGATGATGATGATGATG 1308
      ||||| |
QY 3784 ACCGATGCTCTGGAGAGTGGAGATCTGGACAGGCTCTCTGGGGCAGATGTGTGAT 3843
      ||||| |
Db 1309 GGTCTCTGGCGGAGAGTGGAGATCTCTGACAGGCTCTCTGGGGCAGATGTGTGAT 1368
      ||||| |
QY 3844 GACTCTGGGACCTGGCGGAGGAGTGGTGTGTGATGATGATGATGATGATGATGATGATG 3903
      ||||| |
Db 1369 GACAGCTGGGACCTGGAGAGTGGCGGCTGTGTGTGATGATGATGATGATGATGATGATG 1428
      ||||| |
QY 3904 CTGGCTGCTCTGAGGAGCCTCTGTTTGGCCAGGAGAACTGGAACCATCTGTTGATGAT 3963
      ||||| |
Db 1429 CTCATGCTTGCATTTTGTGGATTTAGTTCAGGATCAGGCGCCATCTGCTGATGATA 1488
      ||||| |
QY 3964 ATGCGGTCAAAGGAATGATTCATTTCTATGGGACTGTACAGCCCAAACTCTGGGAGAC 4023
      ||||| |
Db 1489 CTGAATCTGGGAGGAAGAGTGGCGGCTGTGTGATGATGATGATGATGATGATGATGATG 1548
      ||||| |
QY 4024 AGTGAATGACAGCAAGAGATGCTGGGCTGAGGCTGCTGGA 4068
      ||||| |
Db 1549 CACGACTTGGGATGACAGAGATGCTGGAATCTGCTGCTCAGGA 1593
      ||||| |
```

RESULT 12

SSSRP3

SSSRP3

1665 bp

DNA

linear

MAM 27-AUG-1997

```
DEFINITION S.scrofa mRNA for scavenger-receptor protein, 1665bp.
ACCESSION X99334
VERSION X99334.1 GI:1480360
KEYWORDS WC1 gene.
SOURCE Sus scrofa.
ORGANISM Sus scrofa.
REFERENCE 1 (bases 1 to 1665)
AUTHORS Kanan,J.H., Nayeem,N., Binns,R.M. and Chain,B.M.
TITLE Mechanisms for variability in a member of the scavenger-receptor
cysteine-rich superfamily
JOURNAL Immunogenetics 46 (4), 276-282 (1997)
MEDLINE 97364683
PUBMED 9218529
REFERENCE 2 (bases 1 to 1665)
AUTHORS Kanan,J.H.C.
TITLE Direct Submission
JOURNAL Submitted (15-JUL-1996) J.H.C. Kanan, University College London,
Immunology, Fourth floor, Windeyer Building, Cleveland Street,
London WC1 6BT, UK
FEATURES
source
1..1665
/organism="Sus scrofa"
/strain="SLAB/b"
/db_xref="taxon:9823"
/cell_type="peripheral blood T lymphocyte"
29..1665
/gene="WC1"
29..1665
/gene="WC1"
/codon_start=1
/product="scavenger-receptor protein"
/protein_id="CAA67708.1"
/db_xref="GI:1480361"
/db_xref="SPTREMBL:Q29111"
/translation="MALDRHLSLQGLFLLIVVGGQALELRKDRHRCRGRVELKH
OGKMGTVNAYVGLPFAVVCQLECGTAVDAPVAFYGRGDGPTLWAFEGESTL
TDCRPNITDYRTKGFSDMDAGAVCSGFLVVGDRPCSGRVEHSDEDMTPVSDGN
FTPTAHVICAELGCGKAVSVLGRPLRGPNQVWPEERCKGDEPELWSCPVPDPC
GTCFHSGAVOVCSNAYSEVLRMRNCTSOCAGRVEMNISGGHRTLCASHWTANAVVC
HQLCGVALSTPKNAGDGMKARFHCSGAESFLWNCPTVALGVPDCAHNTASVIC
SGNQTQVLPQCNLSLSEAPGASAEESTNCSESRLRLVDESVCRRVELHOGSW
GTICDSDLDLDAHVCRLQCGDAISALGAFHFGVSGRIWLDVNCVTGKESHLWQC
PFRGWRHDCRHKEDAGVICSLESIALMRVEDQECAGWLEFYNGTWGVCGRSPMEDI
TLSMICRLQCGDSGLTNSVSLRSGSRPRWVDGIQCG"
29..94
sig_peptide
/gene="WC1"
BASE COUNT 344 a 405 c 539 g 377 t
ORIGIN
Query Match 11.9%; Score 551.8; DB 4; Length 1665;
Best Local Similarity 61.6%; Pred. No. 1.4e-126;
Matches 974; Conservative 0; Mismatches 587; Indels 20; Gaps 5;
QY 2086 GATATGGAGCTGAGGCTGTGTGGTGGAGCAGCAGGCTGTGCTGGAAAAGTTGAGGTGAAT 2145
      ||||| |
Db 98 GCTCTGGAGCTGAGCTGAAGCATGGAAGCACCCTGTGTGAGGGAGACTGGAATTTGAAG 157
      ||||| |
QY 2146 GTCCAGGGTCCCTGGGAATTTCTGTGCTAATGCTGGGGGAATGAACATTTGCTGAAGTT 2205
      ||||| |
Db 158 CACCAAGGAAAATGGGGCACTGTGAATCAGGCTTACTGGGGTTTCCAGGAGCAACCGTG 217
      ||||| |
QY 2206 GTTTCAGGCAACTTGAATGTGGTCTGCAATCAGGCTCTCCAGAGAGCTCATTTTCA 2265
      ||||| |
Db 218 GTGTGCAGACAGCTGGAATGTGAACTGCTGTGTATGCTCCCGTGCATCTTTATTTGA 277
      ||||| |
QY 2266 GAAAGAACATTACACATCTTAATGTGCAATTTCTGCTGCACCTGGAG-GGGAAGCCCTCT 2324
      ||||| |
Db 278 CGAGGGGATGGTCCCATCTGTTGTTATGCTTTTGTGAAGGACAGAGTCAACTCTCACT 337
      ||||| |
QY 2325 CTGGGATTTATACATGAGGAGTGGAAACAGACTCGGTGTCATTTTAAATATATGAAGCAG 2384
      ||||| |
```

Db 338 GACTGTAGGCATCTAAATATTACAGACTACAGGACTAAAGGCTTTTCCCATGATTGGGAT 397
QY 2385 TTGTGATCTGCTGAGCCACAGCAGCCAGGCTGGTGGAGCTGATATGCGCTGCTCTGG 2444
Db 398 GCTGGAGAGTCTGCTCAGGATTTGTGC-GCCTGGTGTAGGGAGCAGACCTTGTCTCAGG 456
QY 2445 ACCTGTTGAAGTGAACATATGCAGACACATGGCGCTGTCTGTGATCTGATTTCTCTCT 2504
Db 457 GCGAGTGAAGTGCATCTTGGAGAGACTGGAGCCAGTGTCTGATGGGAACCTTACATT 516
QY 2505 TCATGCTGCCATGCTGCTGAGAGAAATTAATTTGGAGATGCCATCTCTTCTCTG 2564
Db 517 CCCCAGTCCCACTGCTGTCAGAGCTGGGATGGCAAGGCTGTCTGTCTCTGG 576
QY 2565 GGGAGATCACTTTGAAAAGGGAATGCTAACTTGGCGCGAAAGTCCAGTGTGAAGG 2624
Db 577 TCAGAGGCCCTCAGAGAGCCAGGTAACCAAGTCTGGCCTGAAGAGTTCAGGTGTAAAGG 636
QY 2625 GAGTGAACCTCACTTTGCATTATGCCCCATTTGTTCAACATCCGGAAGACACTTGTATCCA 2684
Db 637 GGAGAGCCCTGAGCTGTGCTGCTGCCAGAGTGCCTGTCCAGGAGGCACTTGTCAACCA 696
QY 2685 CACAGAGAGTGGAGTGTCTGTTCCCGATATACAGATGCCGACTTGTG---AATGG 2741
Db 697 CAGTGGAGCTGTTCAAGTGTCTGTTTCAGCATACTCAGAAAGTCCGGCTCATGAGAAATGG 756
QY 2742 CAAATCCAGTGTGACGGGCAAGTGCAGATCAACGTGCTTGGACACTGGGGCTCACTGTG 2801
Db 757 CACTCTCAGTGTGGGGGAGTGGAGATGATATTTCTGGAGGATGGAGACACTCTG 816
QY 2802 TGACACCCACTGGGACCCAGAGATGCCGCTGTTTATGTCAGACACTCAGCTGTGGGAC 2861
Db 817 TGCTCCCACTGACCATGACCAATGCCAATGTGTGTGTCATCAGCTGGCTGTGGAGT 876
QY 2862 TGCTCTCAACACAGGAGGAAATATATTGAGAAAGAGTTCGTGTGTGGGGACA 2921
Db 877 CGCCCTCTCCACCCGAGGGA-----GCAGAGGAGGTGATCAGATGTGGAAGC 927
QY 2922 CAGTTTCATTGTTAGGGAATGAGTCACTTCCTGGATACTGTCAAATGACACTTCTTGG 2981
Db 928 CCGATTTCACTGCTTGGGGCTGAGTCTTCTGTGGAATTTGCTGACTGCCCTGGG 987
QY 2982 AGCACTTCCCTGTATCCATGAATATCTGCTGTGATCTGCACAGAGAGCCCTGACCCA 3041
Db 988 TGTTCCTGACTGTGCCATGAAACACAGCCTCTGTGATCTGCTCAGGAACACAGACCCA 1047
QY 3042 GCACCTGTTTCCATGCTCCCAATGTATCTGACCCATATTTCTGCACTTCCAGAGGG 3101
Db 1048 GGTGCTGCCTCAGTGAATGACTCCCTGTCTGAACCAAGAGGCTCTGCACCCCTCAGAGGA 1107
QY 3102 CAGTGTCTTGTATCTGCTTAGAGGACAAACGGCTCCGCTAGTGGATGGGACAGCCGCTG 3161
Db 1108 AAGCACCACCACTGCTCAGAGAGCAGACAGCTGCGCTGTGTAAGGAAAGTATTG 1167
QY 3162 TGGCGGAGAGTAGAGATCTATCAGACAGCGTCTTGGGGCACCATCTGTGATGACGCTG 3221
Db 1168 CTCACGGAGAGTGGAGATCTCTCACCAGGCTCCTGCGGGCACCATCTGTGATGAAAGCTG 1227
QY 3222 GGACCTGAGCGATCCCACTGGTGTCTCAAAAGCTGGCTGTGGAGTGGCTTCAATGC 3281
Db 1228 GGACCTGGAGATGCGACAGCTGGTGTGTCAGCGAGCTGGGCTGTGGAGATGCCATCAGTGC 1287
QY 3282 CACGGTCTCTGCTCAGTCTGGGAGGGGTGAGGGCCCATCTGGCTGGATGACCTGAAGTGC 3341
Db 1288 ACTGGGTCTGCTCACTTTGGGTGGGCTCAGGAGCAGTCTGGCTGATGACCTGAAGTGC 1347
QY 3342 CACAGGAACGAGTCCCACTTGTGGCAGTGCCTTCCCGGGCTGGGGGGCAGCAGCACTG 3401
Db 1348 TACAGGAAAGAGTCCCACTATGGCAGTGCCTTTTCAGGGGCTGGGGGGGCGCAGCACTG 1407
QY 3402 CAGGCAAGGAGGAGCAGCGGCTCATCTGCTCAGATTCACAGCCTTGGGCTCTACAG 3461
Db 1408 CAGACAAAGAGGAGCAGCGGGTCTATCTGCTCAGAGTCCCTGGCCCTCAGG-----AT 1461

QY 3462 TGAACCTGAAACAGAGAGCTGTGCTGGAGATTTGGAAGTCTTCTATACGGGACCTGGGG 3521
Db 1462 GGTGAGCGAGGAGGAGTGTGCGGGTGGCTGAAGTTTCTTACAAGCGGACCTGGGG 1521
QY 3522 CAGCTCGGCGAGGAGGAACATCACCACAGCAGCCATACAGCAGTGTGTCAGGACGCTGG 3581
Db 1522 CGGTCTGCTGCCACCCCATCAGATCACCTTGTCCATGATCTCGAGCAGCTGG 1581
QY 3582 CTTGGGGAGATGAGTGTGTCAGCTCGCCCTTTATCTAAGACAGGCTCTGTTTCAT 3641
Db 1582 CTTGGGGAGACATGGAACCTCACTTCCGTTCTCTTAGGAGGTTCTAGACCCG 1641
QY 3642 GTGGGTGATGATCAGTTCAAGT 3662
Db 1642 GTGGGTGATGATCAGTTCAAGT 1662

RESULT 13
AC131207/c

LOCUS Homo sapiens clone RP11-326A16, *** SEQUENCING IN PROGRESS ***, 10
DEFINITION AC131207

ACCESSION AC131207

VERSION AC131207.1 GI:22296912

KEYWORDS HTG: HTGS_PHASE1.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 202252)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-oshan, F.R., Allen, C.,
Alsbrooks, S.L., Amarantunge, H.C., Are, J.R., Ayale, M., Banks, I.,
Barbieri, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D.,
Bouch, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Galsi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hoques, M., Holloway, C., Hollins, B.,
Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jollivet, S., Joudah, S.,
Karlssohn, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louisegh, H.,
Lozad, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, B., Mahoney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Ogundun, N., Nickerson, E., Nwokoko, S., Oguh, M., Okwuonu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rivers, M., Rojas, A., Rojokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Schirer, S., Scott, G., Shen, H., Shoshitari, N., Sisson, I.,
Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, R., Thomas, S.,
Tansani, K., Vazquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williams, A., Wlezyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

Direct Submission

Unpublished

TITLE

JOURNAL


```
/codon_start=1
/product="EP4a"
/protein_id="AAK92475.1"
/db_xref="GI:15217047"
/translation="RRSLHCPISLSRQIRLVDGGSGVCRVEILDOGSWGTICD
DQWLDLDARVCRQLGCEALNATGSAHFAGSGGPIWLIDLCNACNEHSHWPCRSRGW
GRHDCRHKEDAGVICSEYLALRMVSEDOECAGLVEFYNGTWGSCVRSRPMDDVTYSII
CSQLGCGSGSLNTSVGLREGSRPHIDLQCRKTDTSLNQPCSPKRFSSCSKPEEA
YISCAQRRKCPAAAPACTDRKRLRGDSECSGRVEVHSGSWGTGCDSDSALAEA
EYVCOGLQGALEAVRAGTDFGNGSIWLDVRCRGRESSLWDCAAEPWGGSDCKHE
EDAGYCVSRVRLTPTTPTAGTINIGLERILRSGLDFGCSLTAQERKPEEPWLVGRS
PGRYVELICLGSWGHFTLMSVQCRI"
BASE COUNT      420 a 566 c 707 g 470 t      2 others
ORIGIN
Query Match      10.3%; Score 476.4; DB:4; Length 2165;
Best Local Similarity 69.5%; Pred. No. 1e-107;
Matches 701; Conservative 1; Mismatches 252; Indels 54; Gaps 2;
QY 3108 TTTGATCTGCTTAGAGACAAACGGCTCCGCTAGTGGATGGGACAGCGCTGTGCCGG 3167
DB 27 TTTCTCTCTCTACAGACAGACAGATCCGCTGTGTGGACGGGGGTTCCTGCGTCGG 86
QY 3168 GAGATAGACATCTATCACGACGCTTCTGGGACCATCTGTGATGACGCTGGACCT 3227
DB 87 GAGATGGAGATCTTGACAGGCTCTCTGGGACCATCTGTGATGACGCTGGACCT 146
QY 3228 GAGCGATGCCACGCTGTGTCAAAAGCTGGCTGTGGAGTGGCTTCAATGCCACGGT 3287
DB 147 GGACGATGCCCGCTGTGTGCAGCAGCTGGCTGTGGAGAGCCCTCAATGCCACGG 206
QY 3288 CTCTGCTCACTTTGGGAGGGGTGAGGGCCATCTGGCTGGATGACCTGAATGACAGG 3347
DB 207 GTCTGCTCACTTGGGAGGATCAGGGCCATCTGGCTGGAGACCTGAATGCGCGCG 266
QY 3348 AACGGAGTCCACTGTGGCAGTCCCTCCCGCGCTGGGGCAGCAGACTGCAGGCA 3407
DB 267 AATGAGTCTCAGTGTGGAGGTGCGCTTCCCGGGCTGGGGCGCAGCAGCTGCAGACA 336
QY 3408 CAAGGAGGAGCGAGGGTCTATCTGCTCAGAAATTCACAGCTTGAGCTCTACAGTAAAC 3467
DB 327 CAAGGAGGAGCGGGGTCTATCTGCTCAGAGTACCTGGCCCTCAGG-----ATGGTGAG 380
QY 3468 TGAACAGAGAGCTGTCTGGGAGATGGAGTCTTCTATACGGGACCTGGGGCAGCGT 3527
DB 381 CGAGGAGGAGGAGTGGCTGGGTGGTGAAGTTTCTACAGGGACCTGGGGCAGTGT 440
QY 3528 CGGAGGAGGAGCAATCACACAGCCATAGCAGGCAATGTGTGAGGAGCTGGGCTGTGG 3587
DB 441 CTCGGCAGCGCCCATGGATGATCTACCGGTGTCCATCATCTGCAGTTCGCTTGGCTGTGG 500
QY 3588 GGAGATGGAGTGTGTCAGGCTCGCCCTTTATCTAAGACAGGCTCTGTTTCTATGTGGGT 3647
DB 501 GGACAGTGAAGTCTCAACACTTCTGTGCTCTCAGGAGAGGTCTAGACCCCGGTGGAT 560
QY 3648 GGATGACATTCAGTCTTAAAGCATATCTCCATATGCACTGCTGTCTGCCCATG 3707
DB 561 AGATTAATTCAGTGTGGAAACTGATACCTCTCTGTGCACTGTCTTCTGGCCCGTG 620
QY 3708 GGAGCAGAAGATCTCCAGCCACAGAGAGAGCTTGATACATGTGAAGATAGAATAA- 3766
DB 621 GAAATTCAGTCTCTCTCCAAAGGAGGAGGCTACATCTCTCTGTGCGAGGAAGAGAC 680
QY 3767 -----GAGTGGCTGGAGG 3779
DB 681 CAAGAGCTGTCCAGCTGCTGCCCTCTGCACAGACAGAGAGAGCTCCGGCTCAGGGGAGG 740
QY 3780 AGACACCGAGTGTCTGGGAGAGTGGAGATCTGGCAGCAGGCTCTCTGGGCGACAGTGTG 3839
DB 741 AGACACCGAGTGTCTAGGCGGGTGGAGGTCTGGCAGACGGGCTCTCTGGGCGACCGTGTG 800
QY 3840 TGATGACTCTCTGGGAGCTGGCGGAGCGGAGGTGTGTGTGACAGCTGGGCTGTGGCTC 3899
|||||
```

```
Db 801 CGATGACTCTCGAGCCTGGCAGAGCGCGAGGTGGTGTGTCTCAGCAGCTGGCTGTGCCCA 860
QY 3900 TGTCTTGGCTGCTCCTGAGGACGCTTCTGTTTGGCCAGGGAACCTGGAACCATCTGTTGGA 3959.
Db 861 GGCCCTTGAAGCGCTGCGGCTGGGACATTCGCGCCCTGGAAATGGGAGCATCTGGCTGGA 920
QY 3960 TGACATCCGCTGCAAGAGAAATCAGTCAATTTCTATGGGACTGTACACCCCAACCTGGG 4019
Db 921 CGAGGTGCGCTGAGGGGCGGGAGTCTCTCTGTTGGGACTGTGCTGCCAGCCCTGGG 980
QY 4020 ACAGATGACTGTGGACACAGGAAGATGTCGCGCTGAGGTGCTGCTGG 4067
Db 981 GCAGCGGACTGCAARACAGGAGGACGCTGGTGTGAGGTGCTCTGG 1028

RESULT 15
SSSRP2
LOCUS      SSSRP2      1316 bp      DNA      linear      MAM 27-AUG-1997
DEFINITION S.scrofa mRNA for scavenger-receptor protein, 1316bp.
ACCESSION X99333
VERSION   X99333.1 GI:1480358
KEYWORDS  WCI gene.
SOURCE    Sus scrofa.
ORGANISM  Sus scrofa.
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
REFERENCE 1 (bases 1 to 1316)
AUTHORS   Kanan, J.H., Nayeem, N., Binns, R.M. and Chain, B.M.
TITLE     Mechanisms for variability in a member of the scavenger-receptor
           cysteine-rich superfamily
JOURNAL   Immunogenetics 46 (4), 276-282 (1997)
MEDLINE   97364683
PUBMED    9218529
REFERENCE 2 (bases 1 to 1316)
AUTHORS   Kanan, J.H.C.
TITLE     Direct Submission
JOURNAL   Submitted (15-JUL-1996) J.H.C.-Kanan, University College London,
           Immunology, Fourth floor, Windeyer Building, Cleveland Street,
           London WC1 6BT, UK

FEATURES             Location/Qualifiers
     source            1..1316
                     /organism="Sus scrofa"
                     /strain="SLAb/b"
                     /db_xref="taxon:9823"
                     /cell_type="peripheral blood T lymphocyte"
                     <1..126
     intron            /number=1
                     join(127..219,301..1316)
     gene              /gene="WCI"
                     join(<127..219,301..>1316)
     CDS                /gene="WCI"
                     /codon_start=3
                     /product="scavenger-receptor protein"
                     /protein_id="CAA67707.1"
                     /db_xref="GI:1480359"
                     /db_xref="SPTREMBL:Q29110"
                     /translation="NQTVLPQCNDSLSEPAASAASECTANCSRQLRLVDGSSYC
                     SGEVILHOGSMGTGCDSDSDLDHVRQVLCGCKTILSALGSAHFAGSGRIWLDDV
                     NGTGKESHLWQCPSRGWOQHNRHKEDAGVICSEYLALRMVSEDOECAGLVEFYNGT
                     WGVCRSPMEDITLITICRLGCGSGTILNSVLRGSRPRWDRIQCRKTDSSLWE
                     CPSDPKYSKSSKEEAYVSCAGRRPKSCPTTAPCTDKELRLRGDSVCSGRVEVWH
                     GGSWCTVCDSDSLAEAEVVCQQLGCGHALEALRDAAAFPGSGSIWLDEVQCRGREGS
                     LWDCAAGPWGQSDCKHEADAGVRCGARPGSNIP"
                     127..219
     exon              /gene="WCI"
                     /number=2
                     220..300
     intron            /gene="WCI"
                     /number=2
                     301..>1316
     exon              /gene="WCI"
                     /number=3
                     264 a 345 c 422 g 285 t

BASE COUNT      264 a 345 c 422 g 285 t
```

Search completed: May 12, 2003, 05:29:51
Job time : 8566 secs

ORIGIN

Query Match 10.0%; Score 460.6; DB 4; Length 1316;
Best Local Similarity 68.8%; Pred. No. 8.7e-104;
Matches 690; Conservative 0; Mismatches 259; Indels 54;

QY	3113	TC	TGCTTTAGAGACAAACGGCTCCGCCCTAGTGGATFGGGACAGCCGCTGTGCGGGAGAG	3172
Db				
Db	292	TC	TATTACACAGCAGACAGACTGCGCCCTGGTGGATGGAGGAGATTACTGCTCAGGAGCG	351
QY	3173	TAG	AGATCTATCACGACGGCTTCTGSGGCACCATCTGTGTATGACGGCTGGGACCTGAGCG	3232
Db				
Db	352	TG	AGATCTCTCACAGGGCTCCTGGGCACCGTCTGTGTATGACAGCTGGACCTGGACG	411
QY	3233	AT	CCCAACGTGGTGTGTCAAAAAGCTGGGCTGTGGAGTGGCCCTTCAATGCCACGGTCTCTG	3292
Db				
Db	412	AT	CGCATGTGGTGTGACGACAGCTAGGCTGTGGAAAAAACCCCTCAGTGCATCGGGTCTG	471
QY	3293	CT	CACCTTTGGGAGGGGTACAGGCCCATCTGGCTGGATGACCTGAACCTGCACAGGAACGG	3352
Db				
Db	472	CT	CACCTTTGGGGGGGGCTCAGACGCATCTGGCTGTGATGTGACCTGAACCTGTACAGGAAAG	531
QY	3353	AG	TCCCACTTGTGGCAGTGCCTTCCCGCGGTGGGGCAGCAGCACTGCAGGCACACAAG	3412
Db				
Db	532	AG	TCCCACTATGGCAGTGCCTTCCCGGGCTGGGGCAGCACAACCTGCAGACACAAG	591
QY	3413	AG	AGCAGGGGTATCTGCTCAGAAATTCACAGCCTTGGAGCTCTCAGTGAACCTGAAA	3472
Db				
Db	592	AG	ACAGGGGTATCTCTCTCAGAGTCCCTGGCCCTCAGG-----ATGGTAGCGCAGG	645
QY	3473	CAG	AGCTGTGCTGGGACATTTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTCGGCA	3532
Db				
Db	646	AC	AGAGTGTGCCGGTGGCTGGAAGTTTCTACACGGGACCTGGGGCGGTGCTGTGCC	705
QY	3533	GG	AGGAACATCACACAGCCATAGCAGGCATTTGTGTSCAGCAGCTGGGCTGTGGGAGA	3592
Db				
Db	706	GC	AGCCCATGGAAGACATCACCTTGTCCATCATCTGCAGGCAGCTGGCTGTGGGACA	765
QY	3593	AT	GGAGTTGTACGCTCGCCCTTTTATCTAAGACAGGCTCTGTGTTTCATGTGGTGTGATG	3652
Db				
Db	766	GT	GGAACCCCTCAACTCTCTGTCTTTCTATTAGGAAGTTCTAGACCCCGGTGGGTGGATA	825
QY	3653	AC	ATTCAAGTGTCTTAAACGCATATCTCCATATGCGAGTGCCTGTCTGCCCATTTGGGAGC	3712
Db				
Db	826	GA	ATCAGTGTGGAACATGATGATCTCTCTGGAGTGTCTTCTCTGACCCCTTGGAAT	885
QY	3713	GA	AGATCTCCAGCCACGACGAAGAGACCTGGATCACATGTGAAGATAGATAA-----	3766
Db				
Db	886	AT	AGTTTCATGTTCTTCAAAGGAGGAAGCCTATGTCTCCTGTGTGAGGAAGAGACCCAAGA	945
QY	3767	-----	-----GAGTCGCTGGAGGAGACA	3784
Db				
Db	946	GC	TGTCCAATCTGCCCCCTGCACAGACAAGAAAGCTGCGACTCAGGGGAGGAGACA	1005
QY	3785	CC	GAGTGTCTTGGGAGATGGAGATCTGCACAGCAGGCTCCTGGGGCAGACAGTGTGTATG	3844
Db				
Db	1006	GC	GTGTGCTCAGGCGCAGTGGAGGTGSCACGGCGGCTCCTGGGCACATCTGTGTGTATG	1065
QY	3845	ACT	CTCTGGGACCTGGCCGAGGGCGGAAGTGTGTGTACAGCAGTGGCTGTGGCTCTGCTC	3904
Db				
Db	1066	ACT	CTCTGGAGCTTGGCAGAGGCGAGGTGTGTGTCTACAGCAGTGGGCTGTGGCCATGCC	1125
QY	3905	TG	CTGCCCTGAGGAGCGTCTGTTTGGCCAGGGAACCTGGAACCATCTGGTTGGATGACA	3964
Db				
Db	1126	TG	AGCCCTTGGGGATGCGGCATTTGGCCCTTGGAAATGGGAGCATCTGGCTTGCACGAGG	1185
QY	3965	TG	CGGTGCAAGGAATAAGTCAATTTCTATTGGACTGTACGCCCAAAACCCCTGGGGACAGA	4024
Db				
Db	1186	TG	CAGTCAGGGGACGGAGTCTCCCTGTGTGGACTGTGCTGTGGGGCCCTGGGGGCGAGA	1245
QY	4025	GT	GACTGTGGACACAAGGAAGATGCTTGGCGTGAGGTGCTCTGG	4067
Db				
Db	1246	GT	ACTCAAGCATGAGGAGGACGCTGGAGTCAGATGCTCTGG	1288

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2003, 02:53:40 ; Search time 3980 seconds

(without alignments)

18832.340 Million cell updates/sec

Title: US-09-759-130b-379

Perfect score: 4628

Sequence: 1 gcgcgcctcgcgatctaga.....aaacggacgcgtggctgcac 4628

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estm:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estl:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	958.2	20.7	999	9	AL529439
2	578.8	12.5	582	12	BG570972
C 3	568.2	12.3	581	14	BQ776070
C 4	512.6	11.1	554	9	AL529438
5	491.4	10.6	940	9	AL550994
C 6	489.4	10.6	491	9	AI095706

C 7	453.8	9.8	457	9	AI738432
C 8	447.4	9.7	449	9	AI609329
C 9	446.2	9.6	451	10	AW082879
C 10	416.8	9.0	799	13	BI185454
C 11	414.4	9.0	429	9	AI278132
C 12	408.4	8.8	718	13	BI183796
C 13	407.8	8.8	411	9	AI090561
C 14	401.4	8.7	483	12	BF132072
C 15	392.8	8.5	606	14	BQ024006
C 16	390.8	8.2	392	14	W42792
C 17	359.4	8.0	371	14	N69291
C 18	359.4	8.0	385	14	W42886
C 19	366	7.9	822	12	BF792807
C 20	353.8	7.9	1100	9	AL540905
C 21	353.4	7.9	1108	9	AL544534
C 22	354	7.6	930	13	BI838644
C 23	324.2	7.0	705	17	BI6308
C 24	320	6.9	320	9	AA582301
C 25	310	6.7	723	13	BI836459
C 26	308.8	6.7	1153	9	AL540904
C 27	306.8	6.6	310	9	AL706449
C 28	304.8	6.6	484	14	N94111
C 29	301.8	6.5	644	13	BI009838
C 30	299.2	6.5	653	10	BE502724
C 31	296.8	6.4	616	13	BI009795
C 32	294.8	6.4	310	9	AI472927
C 33	291.6	6.3	585	14	BQ327934
C 34	280.4	6.3	408	9	AI288981
C 35	289.6	6.3	711	14	BQ378694
C 36	277.4	6.0	583	14	BQ368412
C 37	276.8	6.0	689	12	BF096204
C 38	275.2	5.9	569	10	AW631641
C 39	275.2	5.9	596	13	BI009802
C 40	269.8	5.8	574	14	BQ327932
C 41	268.2	5.8	452	17	AQ672871
C 42	267	5.8	833	9	AL574811
C 43	266.4	5.8	574	13	BI009286
C 44	265.4	5.7	560	13	BI009799
C 45	264.4	5.7	583	14	BQ327909

ALIGNMENTS

RESULT 1
AL529439
LOCUS
DEFINITION AL529439 LTI_NFL001_NBC4 Homo sapiens CDNA clone CS0DD006Y005 5
prime, mRNA sequence.
ACCESSION AL529439
VERSION AL529439.1 GI:12792932
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 999)
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE Full-length CDNA libraries and normalization.
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 999
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DD006Y005"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco kV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850; USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 220 a 236 c 318 g 224 t 1 others
ORIGIN

Query Match 20.7%; Score 958.2; DB 9; Length 999;
Best Local Similarity 99.3%; Pred. No. 6.8e-154;
Matches 982; Conservative 1; Mismatches 4; Indels 2; Gaps 2;

QY 3205 ATCTGTGATGACGGCTGGGACCTGAGCGATGCCAGGTGGTGTGTCTCAAAAGCTGGGCTGT 3264
Db 1 ATCTGTGATGACGGCTGGGACCTGAGCGATGCCAGGTGGTGTGTCTCAAAAGCTGGGCTGT 60

QY 3265 GGAGTGGCTTCAATGCCAGGCTCTGCTCACTTTGGGAGGGTCAAGGCCCATCTGG 3324
Db 61 GGAGTGGCTTCAATGCCAGGCTCTGCTCACTTTGGGAGGGTCAAGGCCCATCTGG 120

QY 3325 CTGGATGACCTGAACCTGCACAGGAGTCCCACTTTGGGAGGTCCCTTCCCGCGGC 3384
Db 121 CTGGATGACCTGAACCTGCACAGGAGTCCCACTTTGGGAGGTCCCTTCCCGCGGC 180

QY 3385 TGGGGGAGCAGCTGACGAGTCAAGAGGAGCGAGGGTCTATCTGTCTCAGAAATTCACA 3444
Db 181 TGGGGGAGCAGCTGACGAGTCAAGAGGAGCGAGGGTCTATCTGTCTCAGAAATTCACA 240

QY 3445 GCCTTGAGGCTTACAGTGAACCTGAACAGAGAGTGTCTGGGAGTGAAGTCTTC 3504
Db 241 GCCTTGAGGCTTACAGTGAACCTGAACAGAGAGTGTCTGGGAGTGAAGTCTTC 300

QY 3505 TATAACGGGACCTGGGACGCTCGGAGGAGAAATCACCACAGCCATAGCAGGCATT 3564
Db 301 TATAACGGGACCTGGGACGCTCGGAGGAGAAATCACCACAGCCATAGCAGGCATT 360

QY 3565 GTGTGAGGAGCTGGGCTGTGGGAGAAATGAGTGTCTAGCCCTCGCCCTTATCTAAG 3624
Db 361 GTGTGAGGAGCTGGGCTGTGGGAGAAATGAGTGTCTAGCCCTCGCCCTTATCTAAG 420

QY 3625 ACAGGCTCTGTTTCATGTGGTGGATGACATTCAGTGTCTTAAACGCATATCTCCATA 3684
Db 421 ACAGGCTCTGTTTCATGTGGTGGATGACATTCAGTGTCTTAAACGCATATCTCCATA 480

QY 3685 TGGCAGTGCCTGTCTGCCCATGGGAGGAGAAATCTCCAGCCACGAGAGACCTGG 3744
Db 481 TGGCAGTGCCTGTCTGCCCATGGGAGGAGAAATCTCCAGCCACGAGAGACCTGG 540

QY 3745 ATCATCTGACAGTGAATAAGAGTCCGAGGAGACACCCAGTGTCTGGGAGAGTG 3804
Db 541 ATCATCTGACAGTGAATAAGAGTCCGAGGAGACACCCAGTGTCTGGGAGAGTG 600

QY 3805 GAGATCTGGACGACGAGCTCTCTGGGACACAGTGTGTGATGACCTCTGGGACCTGGCCGAG 3864
Db 601 GAGATCTGGACGACGAGCTCTCTGGGACACAGTGTGTGATGACCTCTGGGACCTGGCCGAG 660

QY 3865 GCGAAAGTGTGTGACAGAGCTGGGCTGTGGCTCTCTGCTCTGCTGCTGAGGAGCGCT 3924
Db 661 GCGAAAGTGTGTGACAGAGCTGGGCTGTGGCTCTCTGCTCTGCTGCTGAGGAGCGCT 720

QY 3925 TCGTTTGGCCAGGAACTGGAACCATCTGTTGGATGACATGGGTCCAAGGAAATGAG 3984
Db 721 TCGTTTGGCCAGGAACTGGAACCATCTGTTGGATGACATGGGTCCAAGGAAATGAG 780

QY 3985 TCATTTCTATGGGACTGTACAGCCAAACCTCTGGGACAGAGTGAAGTGTGACACAAAGAA 4044
Db 781 TCATTTCTATGGGACTGTACAGCCAAACCTCTGGGACAGAGTGAAGTGTGACACAAAGAA 840

QY 4045 GATCTGGCGTGGAGGTGCTCTGGACAGTCGCTGAATCACTGAATGCTCTCTCAGGTCAAT 4104
Db 841 GATCTGGCGTGGAGGTGCTCTGGACAGTCGCTGAATCACTGAATGCTCTCTCAGGTCAAT 900

QY 4105 TTAGCAGCTATTTTATCCAGTATCTTTGGGCTCTCTCCCTGGTCTCTGTTTATTTCTATTT 4164
Db 901 TTAGCAGCTATTTTATCCAGTATCTTTGGGCTCTCTCCCTGGTCTCTGTTTATTTCTATTT 958

QY 4165 CTCAGCTGGTCCCGAGTTCAGAAACAAA 4193
Db 959 CTCAGCTGGTCCCGAGTTCAGAAACAAA 987

RESULT 2
BG570972 582 bp mRNA linear EST 10-APR-2001
LOCUS 602591546F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4714115 5',
DEFINITION mRNA sequence.
ACCESSION BG570972
VERSION BG570972.1 GI:13578625
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 582)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM1558 row: k column: 12
High quality sequence stop: 576.
Location/Qualifiers

FEATURES

source
1..582
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4714115"
/clone_lib="NIH_MGC_79"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: placenta; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggcgctcgcc); Site_2: SfiI (ggccattatggcc
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-ATTCTAGAGCCGAGCGCGGACATG-dT(30)BN-3',
sequence: 5'-ATTCTAGAGCCGAGCGCGGACATG-dT(30)BN-3',
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."

BASE COUNT 126 a 140 c 198 g 118 t
ORIGIN
Query Match 12.5%; Score 578.8; DB 12; Length 582;
Best Local Similarity 99.7%; Pred. No. 4.6e-89;
Matches 580; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3351 GGAGTCCCCTTGTGGCAGTGCCTTCCCGCGGCTGGGGCAGCAGACTGCAGGCACAA 3410
Db 1 GGAGTCCCCTTGTGGCAGTGCCTTCCCGCGGCTGGGGCAGCAGACTGCAGGCACAA 60

QY 3411 GGAGGACGAGGGGTCTATCTGTCTCAGAAATTCAGCCTTTCAGCTCTACAGTGAAGTGA 3470
Db 61 GGAGGACGAGGGGTCTATCTGTCTCAGAAATTCAGCCTTTCAGCTCTACAGTGAAGTGA 120

QY 3471 AACAGAGAGCTGCTGGGAGATTGGAACTCTTCTATTAACGGGACCTGGGCGACGCTCGG 3530
 Db 121 AACAGAGAGCTGCTGGGAGATTGGAACTCTTCTATTAACGGGACCTGGGCGACGCTCGG 180
 QY 3531 CAGGAGAAATCACCACACGATAGCAGGCAATGTGTGAGGAGCTGGGCTGTGGGA 3590
 Db 181 CAGGAGAAATCACCACACGATAGCAGGCAATGTGTGAGGAGCTGGGCTGTGGGA 240
 QY 3591 GAATGGAGTTGTGAGCTGCTGGGCTTCTATTAAGACAGGCTCTGGTTTCATGTGGGTGA 3650
 Db 241 GAATGGAGTTGTGAGCTGCTGGGCTTCTATTAAGACAGGCTCTGGTTTCATGTGGGTGA 300
 QY 3651 TGACATTCAGTGTCTTAAACAGCATATCTCCATATGSCAGTGGCTGTCTGCCCATGGGA 3710
 Db 301 TGACATTCAGTGTCTTAAACAGCATATCTCCATATGSCAGTGGCTGTCTGCCCATGGGA 360
 QY 3711 GCGAAGAAATCTCCAGCCACAGAGAGACCTGGATCAGATGTAAGATAGAATAGAGT 3770
 Db 361 GCGAAGAAATCTCCAGCCACAGAGAGACCTGGATCAGATGTAAGATAGAATAGAGT 420
 QY 3771 GCGTGGAGGAGACACCGAGTGTCTGGGAGAGTGGAGATCTGGCAGCGAGCTCTGGG 3830
 Db 421 GCGTGGAGGAGACACCGAGTGTCTGGGAGAGTGGAGATCTGGCAGCGAGCTCTGGG 480
 QY 3831 CACAGTGTGTGATGACTCTCTGGGACCTGGCGGAGGCGGAAGTGTGTGAGCAGCTGGG 3890
 Db 481 CACAGTGTGTGATGACTCTCTGGGACCTGGCGGAGGCGGAAGTGTGTGAGCAGCTGGG 540
 QY 3891 CTGTGGCTGTCTGTGCTGCTGGGCTGGGAGCGCTTCTGTGG 3932
 Db 541 CTGTGGCTGTCTGTGCTGCTGGGCTGGGAGCGCTTCTGTGG 582

RESULT 3

BO776070/c
 LOCUS
 DEFINITION UI-H-FHO-bcl-f-19-0-UI.s1 NCI-CGAP_FHO Homo sapiens cDNA clone
 UI-H-FHO-bcl-f-19-0-UI 3', mRNA sequence.

BO776070
 VERSION

BO776070.1 GI:21984546

EST.

SOURCE

ORGANISM

human.
 Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 581)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: James Martin

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained

from Dr. M. Bento Soares, bento-soares@uiowa.edu

The following repetitive elements were found in this cDNA

sequence: 1-39, >AT-richFlow_complexity

Seq primer: M13 FORWARD

POLYA=Yes.

Location/Qualifiers

1. .581

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="UI-H-FHO-bcl-f-19-0-UI"

/clone_lib="NCI-CGAP_FHO"

/tissue_type="Human Chondrosarcoma Cell Line"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

/notes="Organ: Bone; Vector: pT73-Pac (Pharmacia) with a

modified polylinker; Site_1: EcoR I; Site_2: Not I;

RESULT 4

AL529438/c

LOCUS

DEFINITION

AL529438

prime, mRNA sequence.

AL529438

AL529438.1

GI:12792931

KEYWORDS

AL529438 LTI-NFL001-NBC4 Homo sapiens cDNA clone CS0DD006Y005 3
 554 bp mRNA linear EST 13-FEB-2001
 prime, mRNA sequence.
 AL529438
 AL529438.1
 GI:12792931
 EST.

NCI-CGAP_FHO is a cDNA library containing the following tissue(s): Human Grade 1 Chondrosarcoma Cell Line The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dr primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tag for this library is AGAATCCGGC. The cell line was provided by Dr James Martin from University of Iowa
 TAG LIB-UI-H-FHO
 TAG TISSUE=Human Chondrosarcoma Cell Line CS8 - Grade 1 Chondrosarcoma
 TAG_SEQ=AGAATCCGGC"
 BASE COUNT 144 a 125 c 123 g 189 t
 ORIGIN

Query Match 12.3%; Score 568.2; DB 14; Length 581;
 Best Local Similarity 98.6%; Pred. No. 2 9e-87;
 Matches 573; Conservative. 0; Mismatches 8; Indels 0; Gaps 0;

QY 4008 CAAACCCCTGGGACAGAGTGAATGCTGACACAAAGAGATGCTGGCGTGAAGTCTCTGG 4067
 Db 581 CAAACCCCTGGGACAGAGTGAATGCTGACACAAAGAGATGCTGGCGTGAAGTCTCTGG 522
 QY 4068 ACAGTGCCTGAATCAGTGAATGCTGCTCAGTCAATTTAGCAGTATTTATTCAGTAT 4127
 Db 521 ACAGTGCCTGAATCAGTGAATGCTGCTCAGTCAATTTAGCAGTATTTATTCAGTAT 462
 QY 4128 CTTTGGGCTCCTCTCCTGCTGCTGTTTATTCTACGTGGTGGCGAGTTCAGAA 4187
 Db 461 CTTTGGGCTCCTCTCCTGCTGCTGTTTATTCTACGTGGTGGCGAGTTCAGAA 402
 QY 4188 ACAAACATCTGCCCTCAGAGTTTCAACAGAGAGGGGGTCTCTCGAGAGAAATTT 4247
 Db 401 ACAAACATCTGCCCTCAGAGTTTCAACAGAGAGGGGGTCTCTCGAGAGAAATTT 342
 QY 4248 ATTCCATGAGATGGAGACCTGCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4307
 Db 341 ATTCCATGAGATGGAG 282
 QY 4308 TGACACCCCAACCATGTTGTGAAGATGCTAGCGACACATCGCTGTTGGAGTTCCTCC 4367
 Db 281 TGACACCCCAACCATGTTGTGAAGATGCTAGCGACACATCGCTGTTGGAGTTCCTCC 222
 QY 4368 TGCCCTGGAAGCCCAACAAATGACATTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAAT 4427
 Db 221 TGCCCTGGAAGCCCAACAAATGACATTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAAT 162
 QY 4428 ATCTTTGAAGGAGACACAACTTTTAAATGAATAAGAGAGAGAGAGAGAGAGAGAGAGAG 4487
 Db 161 ATCTTTGAAGGAGACACAACTTTTAAATGAATAAGAGAGAGAGAGAGAGAGAGAGAGAG 102
 QY 4488 AAACCTTCCAAATTAACATTTCTTGAACATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4547
 Db 101 AAACCTTCCAAATTAACATTTCTTGAACATAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 42
 QY 4548 TGATAATAAAATTTGAATTTATGTATATCATCTGTTTAAAAAAA 4588
 Db 41 TGATAATAAAATTTGAATTTATGTATATTTAAAAAAA 1

```

SOURCE          human.
ORGANISM        Homo sapiens
REFERENCE       Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheraia; Primates; Catarrhini; Hominiidae; Homo.
TITLE          1 (bases 1 to 554)
JOURNAL        Full-length cDNA libraries and normalization
COMMENT        Unpublished (2001).
                Contact: Genoscope
                Genoscope - Centre National de Sequencage
                BP 191 91006 EVRY cedex - France
                Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES       1..554
                Location/Qualifiers
                source
                1..554
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="CSDD006Y005"
                /clone_lib="LTI_NFL001_NBC4"
                /sex="male"
                /tissue_type="neuroblastoma cells"
                /lab_host="DH10B"
                /note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA
                was primed with a NotI-oligo(dT) primer. Five prime end
                enriched, double-stranded cDNA was digested with Not I and
                cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                vector. Library was normalized. Library was constructed
                by Life Technologies. Contact : Feng Liang Life
                Technologies, a division of Invitrogen 9800 Medical Center
                Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                8371 Email : fliang@lifetech.com URL :
                http://fulllength.invitrogen.com"

BASE COUNT    132 a 166 c 127 g 110 t 19 others
ORIGIN
Query Match    11.1%; Score 512.6; DB 9; Length 554;
Best Local Similarity 94.8%; Pred. No. 9.1e-78;
Matches 525; Conservative 9; Mismatches 19; Indels 1; Gaps 1;

QY 3592 AATGGAGTCTCAGCCTCGCCCTTTA-TCTAAGACAGGCTCTGTTTCATGTGGGTGA 3650
DB 554 AATGGAGTATACGCTCGCCCTTTATCTAAGACAGGCTCTGTTTCATGTGGGTGA 495
QY 3651 TGACATTCAGTCTCTAAACGATATCTCCATATGGCAGTGCCTCTGTCGCCCATGGGA 3710
DB 494 TRACATTCARTGCTCTAAACGATGCTCCATATGGCAGTGCCTCTGTCGCCCATGGGA 435
QY 3711 GCGAAGATCTCAGCCACAGACAGACCTGGATCATGTGAAGATAGATAAGACT 3770
DB 434 GCGAAGATCTCAGCCACAGACAGACCTGGATCATGTGAAGATAGATAAGACT 375
QY 3771 GCGTGGAGGACACACCGAGTCTCTGGAGAGTGGAGATCTGGCAGCGAGGCTCTGGGG 3830
DB 374 GCGTGGAGGACACACCGAGTCTCTGGAGAGTGGAGATCTGGCAGCGAGGCTCTGGGG 315
QY 3831 CACAGTGTGTATGACTCTGGACCTTGGCCGAGCGGAGTGTGTGTCAGCAGCTGGG 3890
DB 314 CACAGTGTGTATGACTCTGGACCTTGGCCGAGCGGAGTGTGTGTCAGCAGCTGGG 255
QY 3891 CTGTGGCTCTGCTCTGCTGCTGCTGAGGACGCTTCTGTTTGGCCAGGAACTGGAACCAT 3950
DB 254 CTGTGGCTCTGCTCTGCTGCTGAGGACGCTTCTGTTTGGCCAGGAACTGGAACCAT 195
QY 3951 CTGTTTGGATGACATGCGGTGCAAGAAATGAGTCAATTTCTATGGAGTGTCAACGCCAA 4010
DB 194 CTGTTTGGATGACATGCGGTGCAAGAAATGAGTCAATTTCTATGGAGTGTCAACGCCAA 135
QY 4011 ACCCTGGGACAGTGTGACACAGGAGATGCTGGCGTAGGTCCTGGACA 4070
DB 134 ACCCTGGGACAGTGTGACACAGGAGTGTGACACAGGAGTGTGACACAGGAGTGTGACAC 75
QY 4071 GTCGCTGAAATCACTGAATGCTCTCAGGTCATTTTACGACTATTTTATCCAGTATCTT 4130
DB 74 GTCGCTGAAATCACTGAATGCTCTCAGGTCATTTTACGACTATTTTATCCAGTATCTT 15

```

```

QY 4131 TGGGCTCCTCTCTCC 4144
DB 14 TGGGCTCCTCTCTCC 1

RESULT 5
AL550994 LTI_NFL006_PL2 Homo sapiens cDNA clone CSODI066YE07 5
LOCUS prime, mRNA sequence.
DEFINITION Full-length cDNA libraries and normalization.
ACCESSION AL550994
VERSION AL550994.1 GI:12888516
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheraia; Primates; Catarrhini; Hominiidae; Homo.
TITLE 1 (bases 1 to 940)
JOURNAL Full-length cDNA libraries and normalization.
COMMENT Unpublished (2001).
                Contact: Genoscope
                Genoscope - Centre National de Sequencage
                BP 191 91006 EVRY cedex - France
                Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES       1..940
                Location/Qualifiers
                source
                1..940
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="CSODI066YE07"
                /clone_lib="LTI_NFL006_PL2"
                /tissue_type="placenta"
                /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                was primed with a NotI-oligo(dT) primer. Five prime end
                enriched, double-stranded cDNA was digested with Not I and
                cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                vector. Library was normalized. Library was constructed by
                Life Technologies. Contact : Feng Liang Life Technologies,
                a division of Invitrogen 9800 Medical Center Drive
                Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                Email : fliang@lifetech.com URL :
                http://fulllength.invitrogen.com"

BASE COUNT    228 a 212 c 279 g 218 t 3 others
ORIGIN
Query Match    10.6%; Score 491.4; DB 9; Length 940;
Best Local Similarity 71.3%; Pred. No. 3e-74;
Matches 645; Conservative 2; Mismatches 258; Indels 0; Gaps 0;

QY 2599 TGGCGCGAAAGTTCCTCAGTGTGAAGGAGTGAACCTCAGCTTGCATTTGCCCATTTGT 2658
DB 9 TGGCGTGAAGATTCAGTGTGAGGAGCATGAGTCCCATCTTCTACTCTGCYCACTAGCA 68
QY 2659 CAACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGATAT 2718
DB 69 CCCCCCCAGAGAGAACCTGTAGCCACAGCAGGAGTGTGGAGTAGTGTCTCAAGATAC 128
QY 2719 ACAGATGTCGCGACTGTGTAATGGCAATCCCAAGTGTGAGCGGCAAGTGGAGATCAACGTG 2778
DB 129 ACAGAAATTCGTTGTTGTAATGGCAAGACCCCGTGTGAGGCGAGAGTGGAGTCAAACG 188
QY 2779 CTTGGACACTGGGGCTCACTGTGTGACACCCACTGGAGCCAGAGAGATCCCGGTGTTCTA 2838
DB 189 CTTGGTCCCTGGGGATCCCTCTGTACTCTCACTGGACATAGAGATGCCCATGTTCTT 248
QY 2839 TGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACACAGGAGGAAATATATTGGAGAA 2898
DB 249 TGCAGACAGCTTAAATGTGGAGTTGCCCTTCTACCCAGGAGGACAGCTTTTGGAAA 308
QY 2899 AGNAGTGTTCGTGTGGGGACACAGGTTTCAATGCTTAGGGAATGAGTCACTTCTGGAT 2958
DB 309 GGAATGGTCAAGATCTGGAGGCATATGTTTCACTGCGACTGGGACGAGCACATGGGA 368

```


Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D.
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 600 Std Error: 0.00
Seq primer: -400p from Gibco.

FEATURES

Location/Qualifiers
1. 457
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2391130"
/clone_lib="NCI_CGAP_Col6"
/tissue_type="colon tumor, RER"
/lab_host="DH10B"
/note="Organ: colon; Vector: pT73b-Pac (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Col6 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1057416-1061255, and 1144584-1145351).
Subtraction by Bento Soares and M. Fatima Bonaudo. "
BASE COUNT 119 a 94 c 100 g 144 t

Query Match 9.8%; Score 453.8; DB 9; Length 457;
Best Local Similarity 99.6%; Pred. No. 1.1e-67;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4128 CTTTGGGCTCTCTCCTGGTCTGTTATTTCTATTTCTACCTGGTCCGAGTTCAGAA 4187
DB 457 CTTTGGGCTCTCTCCTGGTCTGTTATTTCTATTTCTACCTGGTCCGAGTTCAGAA 398
QY 4188 ACAAACATCTGCCCTCAGAGTTTCAACAGAGAGGGGTTCTCTCGAGGAGAATTT 4247
DB 397 ACAAACATCTGCCCTCAGAGTTTCAACAGAGAGGGGTTCTCTCGAGGAGAATTT 338
QY 4248 ATTCCATGATGGAGCTGCTCAAGAGAGAGGACCCACATGGGACAAAGACCTCAGA 4307
DB 337 ATTCCATGATGGAGCTGCTCAAGAGAGAGGACCCACATGGGACAAAGACCTCAGA 278
QY 4308 TGACACCCCAACCATGGTTGTAAGATGCTAGCGACATCGTCTGGGAGTCTTCC 4367
DB 277 TGACACCCCAACCATGGTTGTAAGATGCTAGCGACATCGTCTGGGAGTCTTCC 218
QY 4368 TGCTCTGAAGCCACAAAATGACTTTTAGACTTCCAGGGGCTCACCAGATCAACCTCTAAAT 4427
DB 217 TGCTCTGAAGCCACAAAATGACTTTTAGACTTCCAGGGGCTCACCAGCTCAACCTCTAAAT 158
QY 4428 ATCTTTGAAGGACAGCAACACTTTTAAATGAATAAGAGAGAGTCAAGTTGCCCTATGA 4487
DB 157 ATCTTTGAAGGACAGCAACACTTTTAAATGAATAAGAGAGAGTCAAGTTGCCCTATGA 98
QY 4488 AAACCTGTCCAAATAACATTTCTTTGAACAATAGGACAAAGCTAAATTCATAAAGACTGG 4547
DB 97 AAACCTGTCCAAATAACATTTCTTTGAACAATAGGACAAAGCTAAATTCATAAAGACTGG 38
QY 4548 TGATAATAAAATGAATTATGATATCATCTGTTAAA 4584
DB 37 TGATAATAAAATGAATTATGATATCATCTGTTAAA 1

RESULT 8
AI609329/c 449 bp mRNA linear EST 14-MAY-1999
LOCUS tw30b02.x1 NCI_CGAP_Ov35 Homo sapiens cDNA clone IMAGE:2261163 3',
DEFINITION mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AI609329
AI609329.1 GI:4618496
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 449)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP);
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 631 Std Error: 0.00
Seq primer: -400p from Gibco
High quality sequence stop: 415
POLYA-No.

FEATURES

Location/Qualifiers
1. 449
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2261163"
/clone_lib="NCI_CGAP_Ov35"
/tissue_type="tumor, 5 pooled (see description)"
/lab_host="DH10B"
/note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; This library represents the normalized
version of NCI_CGAP_Ov23. Cloned unidirectionally.
Primer: Oligo dt. Average insert size 0.86 kb. Tumor
types include: mixed Mullerian tumor, papillary serous,
clear cell, spindle cell. All are primary tumors,
metastasis positive. Constructed by Life Technologies."
BASE COUNT 116 a 94 c 97 g 142 t

Query Match 9.7%; Score 447.4; DB 9; Length 449;
Best Local Similarity 99.8%; Pred. No. 1.3e-66;
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4133 GGCTCTCTCTCTGTTCTGTTATTTCTATTTCTACCTGGTCCGAGTTCAGAAACAAA 4192
DB 449 GGCTCTCTCTCTGTTCTGTTATTTCTATTTCTACCTGGTCCGAGTTCAGAAACAAA 390
QY 4193 AACATCTGCCCTCAGAGTTTCAACAGAGAGGGGTTCTCTCAGGAGAGATTTATTC 4252
DB 389 AACATCTGCCCTCAGAGTTTCAACAGAGAGGGGTTCTCTCAGGAGAGATTTATTC 330
QY 4253 ATGAGATGGAGACCTGCCTCAGAGAGAGGACCCACATGGGACAAAGACCTCAGATGACA 4312
DB 329 ATGAGATGGAGACCTGCCTCAGAGAGAGGACCCACATGGGACAAAGACCTCAGATGACA 270
QY 4313 CCCCACCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTGGGAGTCTTCTTCGCT 4372
DB 269 CCCCACCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTGGGAGTCTTCTTCGCT 210
QY 4373 CTGAAGCCCAAAATGACTTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTT 4432
DB 209 CTGAAGCCCAAAATGACTTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTT 150
QY 4433 TGAAGGAGACACAACTTTTAAATGAATAAGAGAGAGTTCAGTTCGCCCTATGGAAACT 4492
DB 149 TGAAGGAGACACAACTTTTAAATGAATAAGAGAGAGTTCAGTTCGCCCTATGGAAACT 90
QY 4493 TGTCCAAATAACATTTCTTTGAACAATAGGAGAACAGCTAAATTTGATAAAGACTGGTGATA 4552

```

|||||
Db 89 TGCCAAATACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATA 30
|||||
QY 4553 ATAAAAATGAATATATATATATCTACTGTT 4581
|||||
Db 29 ATAAAAATGAATATATATATATCTACTGTT 1
|||||

RESULT 9
AW082879/c AW082879 451 bp mRNA linear EST 14-OCT-1999
LOCUS
DEFINITION
  x571g11.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
  IMAGE:2581796 3', mRNA sequence.
ACCESSION
  AW082879
VERSION
  AW082879.1 GI:6038031
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 451)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgaps-r@mail.nih.gov
  This clone is available royalty-free through LLNL; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Seq primer: -40UP from Gibco.
FEATURES
  source
    1..451
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_image="2581796"
    /clone_lib="Soares_NFL_T_GBC_S1"
    /lab_host="DH10B"
    /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
    a modified polylinker; Site 1: Not I; Site 2: Eco RI;
    Equal amounts of plasmid DNA from three normalized
    libraries (fetal lung NBHL19W, testis NHT, and B-cell
    NCI-CGAP-GCBI) were mixed, and ss circles were made in
    vitro. Following HAP purification, this DNA was used as
    tracer in a subtractive hybridization reaction. The driver
    was PCR-amplified cDNAs from pools of 5,000 clones made
    from the same 3 libraries. The pools consisted of
    I.M.A.G.E. clones 297480-302087, 682632-687239,
    726408-728711, and 729096-731399. Subtraction by Bento
    Soares and M. Fatima Bonaldo."
BASE COUNT 115 a 91 c 96 g 149 t
ORIGIN
  Query Match 9.6%; Score 446.2; DB 10; Length 451;
  Best Local Similarity 99.3%; Pred. No. 2.1e-66;
  Matches 448; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4138 CTTCCTCGTGGTCTGTTATTTCTATCTACGTTGCCGAGTTCAGAACAAACAT 4197
|||||
Db 451 CTTCCTCGTGGTCTGTTATTTCTATTTCTACGTTGCCGAGTTCAGAACAAACAT 392
|||||
QY 4198 CTGCCCCCTCAGAGTTTCAACAGAGAGGAGGGTTCTCTCGAGGAGAAATTTATCCATGAG 4257
|||||
Db 391 CTGCCCCCTCAGAGTTTCAACAGAGAGGAGGGTTCTCTCGAGGAGAAATTTATCCATGAG 332
|||||
QY 4258 ATGAGACCTCGCTCAGAGAGAGAGCCACATGGGACAGAACTCAGATGACACCCCC 4317
|||||
Db 331 ATGAGACCTCGCTCAGAGAGAGAGCCACATGGGACAGAACTCAGATGACACCCCC 272
|||||
QY 4318 AACCATGGTCTGAAGTCTAGCGACACATCGCTGTGGGAGTCTCTCGCCTCTGAA 4377
|||||
Db 271 AACCATGGTCTGAAGTCTAGCGACACATCGCTGTGGGAGTCTCTCGCCTCTGAA 212
|||||
QY 4378 GCCAAAAATGACTTTAGACTTCCAGGCTCACCAGATCAACCTCTAAATATCTTTGAAG 4437

```

```

|||||
Db 211 GCCAAAAATGACTTTAGACTTCCAGGCTCCACCAGATCAACCTCTAAATATCTTTGAAG 152
|||||
QY 4438 GAGACAACAACCTTTTAATGAATAAGAGAGAGTCAAGTTGCCCTATGGAAACCTTGTC 4497
|||||
Db 151 GAGACAACAACCTTTTAATGAATAAGAGAGAGTCAAGTTGCCCTATGGAAACCTTGTC 92
|||||
QY 4498 AAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTTGATAAAGACTGGTGATAATAA 4557
|||||
Db 91 AAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTTGATAAAGACTGGTGATAATAA 32
|||||
QY 4558 AATTGAATTATGATATACACTGTTTAAAAAAA 4588
|||||
Db 31 AATTGAATTATGATATACACTGTTTAAAAAAA 1
|||||

RESULT 10
BI185454
LOCUS
DEFINITION
  BI185454 799 bp mRNA linear EST 10-JUL-2001
  UNL-P-FN-cw-d-01-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
  UNL-P-FN-cw-d-01-0-UNL 3', mRNA sequence.
ACCESSION
  BI185454
VERSION
  BI185454.1 GI:14659863
KEYWORDS
  EST.
SOURCE
  Sus scrofa
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
  1 (bases 1 to 799)
  Caetano.A.R., Johnson.R.K. and Pomp.D.
  Generation and sequence characterization of a normalized cDNA
  library from swine ovarian follicles
  Unpublished (2001)
  Contact: Pomp, D
  Department of Animal Science
  University of Nebraska, Lincoln
  Lincoln, NE 68583-0908, USA
  Tel: 402 472 6416
  Fax: 402 472 6362
  Email: dpomp@unl.edu
  Oligo-dt track not found, Not I site shown in beginning of sequence
  is likely internal to the message.
  Seq primer: M13 -29
  POLYA-No.
FEATURES
  Location/Qualifiers
    1..799
    /organism="Sus scrofa"
    /strain="University of Nebraska, Lincoln Swine Selection
    Lines"
    /db_xref="taxon:9823"
    /clone="UNL-P-FN-cw-d-01-0-UNL"
    /clone_lib="UNL-P-FN"
    /dev_stage="ADULT"
    /lab_host="DH10B (Life Technologies)"
    /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
    polylinker; Site 1: Not I; Site 2: Eco RI; The UNL-P-FN
    library is a normalized library representing porcine
    ovarian follicles, ranging between 2.0 to 10.0 mm in
    diameter, collected during 7 days of the follicular phase
    of the pig estrous cycle. This library was derived from
    the library UNL-P-F2. The tag is a string of 5-6
    nucleotides present between the Not I site and the
    oligo-dt track. The library was constructed as described
    by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
    , 1996."
BASE COUNT 206 a 174 c 177 t 1 others
ORIGIN
  Query Match 9.0%; Score 416.8; DB 13; Length 799;
  Best Local Similarity 71.1%; Pred. No. 1.7e-61;
  Matches 564; Conservative 0; Mismatches 228; Indels 1; Gaps 1;

```


Db	312	GAGAGACCTGTGCAGGCGCCTGGNAGCTTTTTCACACGGAGCTTGGGGCAGCGCTTGGCA	371
QY	3533	GGAGGAACATCACACACAGCCATAGCAGGCATTTGTGTGCAGCAGCTGGCTGGTGGGGAGA	3592
Db	372	AGAATAGCATGTCTCCAGCCACAGCTGGGGTGGTATGCAGCAGCTGGCTGTGCAGACA	431
QY	3593	ATGGAGTTGTTCAGCTCGCCCTTTTATCTTAAGACAGGCTCTGGTTCTCATGTGGGTGGATG	3652
Db	432	GAGGGGACATCAGCCCTGTCATCTTCAGACAAGACAGTGTCCAGGCACATGTGGTGGACA	491
QY	3653	ACATTAGCTCTCTAAACCATATCTCCATATATGCGAGTGCCTCTCTGCCCATGGGAGC	3712
Db	492	ATGTTTCTAGTCTCTAAAGGACCTGACACCCCTATGGCAGTGGCCATCATCTCCATGGAAGA	551
QY	3713	GAAGAATCTCCAGCCAGCAGAGACACCTGGATTCACATGTGAAGATAGAATAAGAGTGC	3772
Db	552	AGAGCTGGCAGCCCTCAGAGGAGACATGGATTCACATGTGCCACAAAATAAGACTTC	611
QY	3773	GTGGAGGAGACAGGATGTCTCTGGGAGTGGAGATCTTGGCAGCAGGCTCTCTGGGGCA	3832
Db	612	AAGAAGNAAACATAATTTGTTCTTNGAGCTGTGGAGATCTGTTAGCGNGGTTCTCTGGGGCA	671
QY	3833	CAGTGTCTGATGACTCTCTGGGACCTG 3858	
Db	672	CTGTGTGTGACGACTCTCTGGGACTTG 697	
RESULT	13		
LOCUS	AI090561/c		
DEFINITION	AI090561 411 bp mRNA linear EST 18-AUG-1998 q646907.x1 Soares_fetal_heart_NBHH19W Homo sapiens cDNA clone IMAGE:1691580 3', mRNA sequence.		
ACCESSION	AI090561		
VERSION	AI090561.1	GI:3429620	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 411)		
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .		
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
JOURNAL	Unpublished (1997)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapsb@email.nih.gov This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: -40ml3 fwd. ET from Amersham.		
FEATURES	Location/Qualifiers		
source	1..411 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:1691580" /clone_lib="Soares_fetal_heart_NBHH19W" /sex="unknown" /dev_stage="19 weeks" /lab_host="DH10B (ampicillin resistant)" /notes="Organ: heart; Vector: pF7m3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTACCAATCTCAAGTGGAGCGCGCATCTTTTTTTTTTTT 3'] double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBHH19W."		
BASE COUNT	94 a	81 c	84 g 152 t
ORIGIN			

Query Match 8.8%; Score 407.8; DB 9; Length 411;
 Best Local Similarity 99.5%; Pred. No. 7.8e-60;
 Matches 409; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4192 AAACATCTCCCTCAGAGTTTCAACAGAGGAGGGGTTCTCTCGAGGAGATTATTC 4251
 DB 411 AAACATCTCCCTCAGAGTTTCAACAGAGGAGGGGTTCTCTCGAGGAGATTATTC 352
 QY 4252 CATGAGATGGAGACCTGCTCAAGAGAGAGAGGCCACATGGGACCAAGAACCTCAGATGAC 4311
 DB 351 CATGAGATGGAGACCTGCTCAAGAGAGAGAGGCCACATGGGACCAAGAACCTCAGATGAC 292
 QY 4312 ACCCCCAACCATGGTTGCAAGATGCTAGGACACATGCTGTTGGAGTTCTTCTGCGC 4371
 DB 291 ACCCCCAACCATGGTTGCAAGATGCTAGGACACATGCTGTTGGAGTTCTTCTGCGC 232
 QY 4372 TCTGAAGCCACAAATGACTTTAGACTTCCAGGGTCCACGAGTCAACCTCTAAATATCT 4431
 DB 231 TCTGAAGCCACAAATGACTTTAGACTTCCAGGGTCCACGAGTCAACCTCTAAATATCT 172
 QY 4432 TTGAAGGAGACAACTTTTAAATGAATAAGAGAGAGTCAAGTTGCCCTATGGAACAC 4491
 DB 171 TTGAAGGAGACAACTTTTAAATGAATAAGAGAGAGTCAAGTTGCCCTATGGAACAC 112
 QY 4492 TTGTCACAAATACATTTCTTGACATAGGACACAGCTAAATTTGATAAGACTGGTGAT 4551
 DB 111 TTGTCACAAATACATTTCTTGACATAGGACACAGCTAAATTTGATAAGACTGGTGAT 52
 QY 4552 AATAAAATTTGAATTATGATATCTACCTGTTAAAAAATTTAAAAAATTTAAAAA 4602
 DB 51 AATAAAATTTGAATTATGATATCTACCTGTTAAAAAATTTAAAAAATTTAAAAA 1

RESULT 14
 BFI32072
 LOCUS 601821079F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:4052947 5',
 DEFINITION mRNA sequence.

ACCESSION BFI32072
 VERSION BFI32072.1 GI:10971112
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC <http://mhc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC

CDNA Library Preparation: CLONETECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCW890 row: n column: 20
 High quality sequence stop: 483.
 Location/Qualifiers

FEATURES
 source 1..483
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4052947"
 /clone_lib="NIH_MGC_58"
 /tissue_type="hypertrophoma"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="organ: kidney; Vector: pDNR-LIB (Clontech); Site_1:
 SfiI (ggcccattggcc); Site_2: SfiI (ggcccattggcc);
 Double-stranded cDNA was prepared from cell line RNA.
 and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CAGGGCCATTATGGCC-3' and 3' adaptor

sequence: 5'-ATTCTAGAGCCGAGGCGCGGACATG-dt(30)BN-3'
 (where B = A, C, or G and N = A, C, G, or T). Average
 insert size 1.35 kb (range 0.9-4.0 kb). 15/15 colonies
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA).
 BASE COUNT 136 a 115 c 106 g 126 t
 ORIGIN

Query Match 8.7%; Score 401.4; DB 12; Length 483;
 Best Local Similarity 93.6%; Pred. No. 9e-59;
 Matches 452; Conservative 0; Mismatches 1; Indels 30; Gaps 2;

QY 4050 TGGCGTAGGTCCTCTGGACAGTCGTGAATCACTGAATGCTCTCAGGTCATTATAGC 4109
 DB 1 TGGCGTAGGTCCTCTGGACAGTCGTGAATCACTGAATGCTCTCAGGTCATTATAGC 60

QY 4110 ACTTATTTTATCCAGTATCTTTGGGCTCTCTCTCGTGTGTTTATTTCTTCTCAC 4169
 DB 61 ACTTATTTTATCCAGTATCTTTGGGCTCTCTCTCGTGTGTTTATTTCTTCTCAC 120

QY 4170 GTGTGCGGAGTTCAGAAACAAAACATCTGCCCTTCAGAG----- 4210
 DB 121 GTGTGCGGAGTTCAGAAACAAAACATCTGCCCTTCAGAGTTCAGAGCCCACTAGGAT 180

QY 4211 -----TTTCAACCAAGAGGAGGGTCTCTCGAGGAGAAATTATTCATGAGATG 4260
 DB 181 TGTATAAAATTTTCAACCAAGAGGAGGGTCTCTCGAGGAGAAATTATTCATGAGATG 240

QY 4261 GAGACTGCTCAAGAGAGAGGACCCACATGGACAAAGAACCTCAGATGACACCCCAAC 4320
 DB 241 GAGACTGCTCAAGAGAGAGGACCCACATGGACAAAGAACCTCAGATGACACCCCAAC 300

QY 4321 CATGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCTCGCTCTGAGGCC 4380
 DB 301 CATGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTTCTTCTCGCTCTGAGGCC 360

QY 4381 ACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGAGGAG 4440
 DB 361 ACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGAGGAG 420

QY 4441 ACAACAACCTTTAAATGAATAAA-GAGGAAGTCAAGTTCCTCTGAAACCTTGTCCTCAA 4499
 DB 421 ACAACAACCTTTAAATGAATAAAACGAGGAAGTCAAGTTCCTCTGAAACCTTGTCCTCAA 480

QY 4500 ATA 4502
 DB 481 ATA 483

RESULT 15
 BQ024006/c
 LOCUS BQ024006
 DEFINITION UI-1-BB1p-auu-d-07-0-UI.s1 NCI-CGAP.P16 Homo sapiens cDNA clone
 UI-1-BB1p-auu-d-07-0-UI 3', mRNA sequence.

ACCESSION BQ024006
 VERSION BQ024006.1 GI:19759285
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 606)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Steven Brown
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

THIS PAGE BLANK (USPTO)

GanCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2003, 02:56:30 ; Search time 614 Seconds
(without alignments)
9366.838 Million cell updates/sec

Title: US-09-759-130B-379

Perfect score: 4628

Sequence: 1 gggccgctcgatctaga.....aaacgacgcgtgggtcgac 4628

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 783854 seqs, 621352466 residues

Total number of hits satisfying chosen parameters: 1567708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4628	100.0	4628	9	US-10-042-431-9
2	4628	100.0	4628	9	US-09-759-130B-379
3	4359	94.2	4359	9	US-10-042-431-10
4	4359	94.2	4359	9	US-09-759-130B-380
5	886.8	19.2	4308	9	US-10-042-431-79
6	886.8	19.2	4308	9	US-09-759-130B-449
7	447.4	9.7	449	10	US-09-867-701-10340
8	434.6	9.4	473	9	US-09-918-995-32438
9	213	4.6	599	9	US-10-040-739-138
10	199.2	4.3	461	9	US-09-796-692-269
11	199.2	4.3	461	9	US-09-796-692-7354
12	199.2	4.3	461	9	US-10-040-862-269
13	199.2	4.3	461	9	US-10-040-862-7354
14	197.6	4.3	2562	9	US-09-147-947-5
15	193.2	4.2	2614	9	US-09-147-947-3
16	173.4	3.7	444	10	US-09-925-301-696
17	169.8	3.7	438	9	US-09-918-995-35432
18	149.6	3.2	4344	10	US-09-917-800A-1712
19	143.2	3.1	697	10	US-09-925-297-237

20	134.4	2.9	713	9	US-10-040-739-1356	Sequence 1356, App
21	123.4	2.7	1686	9	US-09-905-291A-147	Sequence 147, App
22	123.4	2.7	1686	9	US-09-902-853-147	Sequence 147, App
23	123.4	2.7	1686	9	US-09-907-824-147	Sequence 147, App
24	123.4	2.7	1686	9	US-09-907-841-147	Sequence 147, App
25	123.4	2.7	1686	9	US-09-904-011-147	Sequence 147, App
26	123.4	2.7	1686	9	US-09-906-742-147	Sequence 147, App
27	123.4	2.7	1686	9	US-09-906-838-147	Sequence 147, App
28	123.4	2.7	1686	9	US-09-907-613-147	Sequence 147, App
29	123.4	2.7	1686	9	US-09-907-942-147	Sequence 147, App
30	123.4	2.7	1686	9	US-10-227-884-113	Sequence 113, App
31	123.4	2.7	1686	9	US-09-904-820-147	Sequence 147, App
32	123.4	2.7	1686	9	US-09-904-859-147	Sequence 147, App
33	123.4	2.7	1686	9	US-09-909-204-147	Sequence 147, App
34	123.4	2.7	1686	9	US-10-230-163-113	Sequence 113, App
35	123.4	2.7	1686	9	US-09-904-786-147	Sequence 147, App
36	123.4	2.7	1686	9	US-09-906-646-147	Sequence 147, App
37	123.4	2.7	1686	9	US-09-906-700-147	Sequence 147, App
38	123.4	2.7	1686	9	US-09-902-903-147	Sequence 147, App
39	123.4	2.7	1686	9	US-09-903-749A-147	Sequence 147, App
40	123.4	2.7	1686	9	US-09-903-786-147	Sequence 147, App
41	123.4	2.7	1686	9	US-10-218-631-113	Sequence 113, App
42	123.4	2.7	1686	9	US-10-230-338-113	Sequence 113, App
43	123.4	2.7	1686	9	US-09-902-736-147	Sequence 147, App
44	123.4	2.7	1686	9	US-09-904-119-147	Sequence 147, App
45	123.4	2.7	1686	9	US-09-904-956-147	Sequence 147, App

ALIGNMENTS

RESULT 1

US-10-042-431-9
; Sequence 9, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; FILE REFERENCE: 10147-602
; CURRENT APPLICATION NUMBER: US/10/042.431
; PRIOR FILING DATE: 2001-10-25
; PRIOR FILING DATE: 1999-06-14
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 4628
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-431-9

Query Match	100.0%	Score 4628	DB 9	Length 4628	
Best Local Similarity	100.0%	Pred. No. 0			
Matches 4628	Conservative	0	Mismatches	0	Gaps
0					
Qy	1	CGGCGCGCTCGGATCTAGAACTAGTATGATGCTGCTCAAAACTCTGGCATATTGAT	60		
Db	1	CGGCGCGCTCGGATCTAGAACTAGTATGATGCTGCTCAAAACTCTGGCATATTGAT	60		
Qy	61	TTTGAAGATGCTGCTGTCATCAGAACCTTTTCTCTGCTGGTAACCTGCTGCTC	120		
Db	61	TTTGAAGATGCTGCTGTCATCAGAACCTTTTCTCTGCTGGTAACCTGCTGCTC	120		
Qy	121	CTGAATTCCTCTTCTCATCAGAGTTTTTAATGAACAGATTTGGAGTTGAGGCTGTC	180		
Db	121	CTGAATTCCTCTTCTCATCAGAGTTTTTAATGAACAGATTTGGAGTTGAGGCTGTC	180		

QY 181 AATGGACGGTCCCTGCTCTGGACAGTGGAGTGAATTCACGGACAGTGGGGACT 240
DB 181 AATGGACGGTCCCTGCTCTGGACAGTGGAGTGAATTCACGGACAGTGGGGACT 240
QY 241 GTGTGTGATGATGGGTGGACACTACTGCTCAACTGCTGTGTCGAACAGCTTGGATGT 300
DB 241 GTGTGTGATGATGGGTGGACACTACTGCTCAACTGCTGTGTCGAACAGCTTGGATGT 300
QY 301 CCAATTTCTTTCCCATGTTTCGTTTGGACAAGCCGCTGACTAGACATGGAATAATTTGG 360
DB 301 CCAATTTCTTTCCCATGTTTCGTTTGGACAAGCCGCTGACTAGACATGGAATAATTTGG 360
QY 361 CTTGATGATGTTTCCCTGTTATGGAATGAGTCAAGCTCTCTGGGAATGTCACACCCGGAA 420
DB 361 CTTGATGATGTTTCCCTGTTATGGAATGAGTCAAGCTCTCTGGGAATGTCACACCCGGAA 420
QY 421 TGGGGAAGCCATAACTGTTTATCATGAGAAGATGTTGGTGTGNACTGTTATGGTGAAGCC 480
DB 421 TGGGGAAGCCATAACTGTTTATCATGAGAAGATGTTGGTGTGNACTGTTATGGTGAAGCC 480
QY 481 AATCTGGGTTGAGGCTAGTGGATGGAACAACCTCCTGTTCAAGGAGAGTGGAGGTGAAA 540
DB 481 AATCTGGGTTGAGGCTAGTGGATGGAACAACCTCCTGTTCAAGGAGAGTGGAGGTGAAA 540
QY 541 TTCCAAGAAAGTGGGGGACTATATGTATGATGATGGGTGGAACCTTGAATACTGTCGCCGTG 600
DB 541 TTCCAAGAAAGTGGGGGACTATATGTATGATGATGGGTGGAACCTTGAATACTGTCGCCGTG 600
QY 601 GTGTGACGCAACTAGGATGCTCATCTCTTTTATTTCTCTGAGCTGTTTAAATACCCCT 660
DB 601 GTGTGACGCAACTAGGATGCTCATCTCTTTTATTTCTCTGAGCTGTTTAAATACCCCT 660
QY 661 GCTGTATTCGCCGCCATTTGGCTGSGATGACATTTTATGCCAGGGGAATGAGTTGGCACTC 720
DB 661 GCTGTATTCGCCGCCATTTGGCTGSGATGACATTTTATGCCAGGGGAATGAGTTGGCACTC 720
QY 721 TGAANTGCAGACATCGTGGATGGGAATCATGACTGACGTACACATGAGGTGTCACA 780
DB 721 TGAANTGCAGACATCGTGGATGGGAATCATGACTGACGTACACATGAGGTGTCACA 780
QY 781 TTAACCTGTTATGATAGTAGTATGTTGAACCTGTAAGCTTGTAGTGGAACTAACCCCTGT 840
DB 781 TTAACCTGTTATGATAGTAGTATGTTGAACCTGTAAGCTTGTAGTGGAACTAACCCCTGT 840
QY 841 ATGGGAGAGTAGAGTGAATAATCCAAGGAAGTGGGGACCGTATGCCACCAATAAGTGG 900
DB 841 ATGGGAGAGTAGAGTGAATAATCCAAGGAAGTGGGGACCGTATGCCACCAATAAGTGG 900
QY 901 AACAACTGCTCAGCTGATGTCGTATGCAAGCAGTTGGGATGTGGAACCCGACTTCACCTTC 960
DB 901 AACAACTGCTCAGCTGATGTCGTATGCAAGCAGTTGGGATGTGGAACCCGACTTCACCTTC 960
QY 961 GCTGGCTTGCTCATTTGCACTCAGGCTGATGTTGTATGGCTTGATGGTGTCTCCTGC 1020
DB 961 GCTGGCTTGCTCATTTGCACTCAGGCTGATGTTGTATGGCTTGATGGTGTCTCCTGC 1020
QY 1021 TCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCGGAACCCGCTCAATTTGACTGT 1080
DB 1021 TCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCGGAACCCGCTCAATTTGACTGT 1080
QY 1081 CTTTCATCAAAACCATGTTCTGTGATCTGCTCAGATGGAGCAGATTTGGNACTGGCACA 1140
DB 1081 CTTTCATCAAAACCATGTTCTGTGATCTGCTCAGATGGAGCAGATTTGGNACTGGCACA 1140
QY 1141 GCAGATGGAAGTAAACATTTCTCAGGAGAGTAGAGGTGAGAAATTCATGAACAGTGGTG 1200
DB 1141 GCAGATGGAAGTAAACATTTCTCAGGAGAGTAGAGGTGAGAAATTCATGAACAGTGGTG 1200
QY 1201 ACAATATGTGACCAGAACTGGAGAGATGAACAAGCCCTTGTGTTGTGAAGCAGCTAGGA 1260
DB 1201 ACAATATGTGACCAGAACTGGAGAGATGAACAAGCCCTTGTGTTGTGAAGCAGCTAGGA 1260
QY 1261 TGTCCGTTTCAGCGCTCTTTGGCAGTCGTCGTGTTAAACCTAGTAATGAGACTAGACACT 1320

DB 1261 TGTCCGTTTCAGCGCTCTTTGGCAGTCGTCGTGTTAAACCTAGTAATGAGCTAGACACT 1320
QY 1321 TCGATAAACAAGCATATCTTGGCACTGGGAATGAGTCAAGCTCTCTGGACTGCACATATGAT 1380
DB 1321 TCGATAAACAAGCATATCTTGGCACTGGGAATGAGTCAAGCTCTCTGGACTGCACATATGAT 1380
QY 1381 GGAAGAGCAAGCAACATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTTCTGATAAG 1440
DB 1381 GGAAGAGCAAGCAACATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTTCTGATAAG 1440
QY 1441 GCAGATCTGGACCTTAAGGCTTGTGGGGCTCATAGCCCTGTTATGSGAGATGGAGGTG 1500
DB 1441 GCAGATCTGGACCTTAAGGCTTGTGGGGCTCATAGCCCTGTTATGSGAGATGGAGGTG 1500
QY 1501 AATATCCAAGAGAGTGGGGGACTGTGTCTCATGACAGATGAGCAACAAGGAATGCAGCT 1560
DB 1501 AATATCCAAGAGAGTGGGGGACTGTGTCTCATGACAGATGAGCAACAAGGAATGCAGCT 1560
QY 1561 GTTGTGTGTAACAATTTGGGATGTGGAAGCCCTATGCAATGTTTGGTATGACCTATTTT 1620
DB 1561 GTTGTGTGTAACAATTTGGGATGTGGAAGCCCTATGCAATGTTTGGTATGACCTATTTT 1620
QY 1621 AAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTCTTGCATTTGCAATGAGTCAAT 1680
DB 1621 AAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTCTTGCATTTGCAATGAGTCAAT 1680
QY 1681 ATCTGGGACTGTGAACACAGTGGATGGGAAAGCAATAATTTGTACACAGAGAGGATGTG 1740
DB 1681 ATCTGGGACTGTGAACACAGTGGATGGGAAAGCAATAATTTGTACACAGAGAGGATGTG 1740
QY 1741 ATTGTAACCTGCTCAGTGTATGCAACATGGGCTGAGGCTGGTGGGCGGCAACCCG 1800
DB 1741 ATTGTAACCTGCTCAGTGTATGCAACATGGGCTGAGGCTGGTGGGCGGCAACCCG 1800
QY 1801 TCTCGGGAGAGTGGAGGTGACTTTCAAGGACGCTGGGACACAGTGTGATGACGGC 1860
DB 1801 TCTCGGGAGAGTGGAGGTGACTTTCAAGGACGCTGGGACACAGTGTGATGACGGC 1860
QY 1861 TGAACAGTAAAGCTGCAGCTGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCAT 1920
DB 1861 TGAACAGTAAAGCTGCAGCTGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCAT 1920
QY 1921 GGCATGGGCTGGGAACCCCTCTACAGATATGGAATAATTTGGCTCGATGATTTCC 1980
DB 1921 GGCATGGGCTGGGAACCCCTCTACAGATATGGAATAATTTGGCTCGATGATTTCC 1980
QY 1981 TGTGATGGAGATGAGTCAAGTCTCTGTCATGCAGCAACAGTGGTGGGAAATAATGAC 2040
DB 1981 TGTGATGGAGATGAGTCAAGTCTCTGTCATGCAGCAACAGTGGTGGGAAATAATGAC 2040
QY 2041 TGCAGTCAAGTGAAGATTTGGAGTGTCTGTTGATGTCATCGGATATGGAGCTGAGG 2100
DB 2041 TGCAGTCAAGTGAAGATTTGGAGTGTCTGTTGATGTCATCGGATATGGAGCTGAGG 2100
QY 2101 CTTTGGGTTGGAGACAGCAGTGTGCTGGAAGTTGAGGTGATGTCAGGGTGGCGGTG 2160
DB 2101 CTTTGGGTTGGAGACAGCAGTGTGCTGGAAGTTGAGGTGATGTCAGGGTGGCGGTG 2160
QY 2161 GGAATTTCTGTGCTAAATGGCTGGGAATGAACATTTGCTGAAGTTGTTGCAGCAACT 2220
DB 2161 GGAATTTCTGTGCTAAATGGCTGGGAATGAACATTTGCTGAAGTTGTTGCAGCAACT 2220
QY 2221 GAATGTGGTCTGCAATCAGGGTCTCCAGAGAGCCCTCATTTTCAGAGAAACATTTACAC 2280
DB 2221 GAATGTGGTCTGCAATCAGGGTCTCCAGAGAGCCCTCATTTTCAGAGAAACATTTACAC 2280
QY 2281 ATCTTAATGCTGAATTTCTGGCTGCACTGGAGGAGCCCTCTCTCTGGGATTTATACGA 2340
DB 2281 ATCTTAATGCTGAATTTCTGGCTGCACTGGAGGAGCCCTCTCTCTGGGATTTATACGA 2340
QY 2341 TGGAGTGGAAACAGACTGCTGTCATTTAAATATGGAAGCAAGTTTGTCTCTCAGCC 2400

Db 2341 TGGAGTGGAAACAGAGCTGCTGTCATTTAAATATGGAAGCAAGTTGATCTGCTCAGCC 2400
Qy 2401 CACAGGAGCCAGCGCTGGTGGAGCTCATATGCCCTGCTCTGACAGTGTGAAAGTAAA 2460
Db 2401 CACAGGAGCCAGCGCTGGTGGAGCTGATATGCCCTGCTCTGACAGTGTGAAAGTAAA 2460
Qy 2461 CATGACAGACATGGCGCTGCTGTGTGATGATGCTGATGCTGATGCTGATGCTGATGCTG 2520
Db 2461 CATGACAGACATGGCGCTGCTGTGTGATGATGCTGATGCTGATGCTGATGCTGATGCTG 2520
Qy 2521 CTGTGACAGAAATTAATTTGGAGATGCCATATCTCTTCTGTGGAGATCACTTTGGA 2580
Db 2521 CTGTGACAGAAATTAATTTGGAGATGCCATATCTCTTCTGTGGAGATCACTTTGGA 2580
Qy 2581 AAAGGGAATGCTTAACCTTGGCCCGAAAGTTCCAGTGTGAAGGAGTGAACACTCACCTT 2640
Db 2581 AAAGGGAATGCTTAACCTTGGCCCGAAAGTTCCAGTGTGAAGGAGTGAACACTCACCTT 2640
Qy 2641 GCATTAATGCCCATTTGTTCAACATCCCGAAGACACTTGTATCCACAGCAGAGAAAGTTGGA 2700
Db 2641 GCATTAATGCCCATTTGTTCAACATCCCGAAGACACTTGTATCCACAGCAGAGAAAGTTGGA 2700
Qy 2701 GTTGTCTGTTCCCATATACAGATGTCGGACTTGTGAATGGCAATCCAGTGTGACGGG 2760
Db 2701 GTTGTCTGTTCCCATATACAGATGTCGGACTTGTGAATGGCAATCCAGTGTGACGGG 2760
Qy 2761 CAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACGTGTGTGACACCCACTGGGACCCA 2820
Db 2761 CAAGTGGAGATCAACGTGCTTGGACACTGGGGCTCACGTGTGTGACACCCACTGGGACCCA 2820
Qy 2821 GAAGATGCCGTGTTCTATGACAGACAGCTCAGCTGTGGAGTGTCTCTCAACACACAGGA 2880
Db 2821 GAAGATGCCGTGTTCTATGACAGACAGCTCAGCTGTGGAGTGTCTCTCAACACACAGGA 2880
Qy 2881 GGAAATATATGGAGAAAGTGTGCTGTGTGGGACACAGGTTTCATTTGCTTAGGG 2940
Db 2881 GGAAATATATGGAGAAAGTGTGCTGTGTGGGACACAGGTTTCATTTGCTTAGGG 2940
Qy 2941 AATGAGTCACTTCCTGGATACTGCAATGACAGTTCCTGGAGCACCTCCCTGTATCCAT 3000
Db 2941 AATGAGTCACTTCCTGGATACTGCAATGACAGTTCCTGGAGCACCTCCCTGTATCCAT 3000
Qy 3001 GGAATACTGCTCTGTGATCTGCACAGAGAAAGCTGACCCAGCCACTGTTTCCATGCGCTC 3060
Db 3001 GGAATACTGCTCTGTGATCTGCACAGAGAAAGCTGACCCAGCCACTGTTTCCATGCGCTC 3060
Qy 3061 GCAATGTATCTGACCCATATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
Db 3061 GCAATGTATCTGACCCATATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
Qy 3121 GAGGACAAACGGCTCCGCTAGTGGATGGGACAGCCGCTGTGCGGGAGAGTAGAGATC 3180
Db 3121 GAGGACAAACGGCTCCGCTAGTGGATGGGACAGCCGCTGTGCGGGAGAGTAGAGATC 3180
Qy 3181 TATCAGCAGCGCTTCTGGGGACCATCTGTGATGACGCTGGGACCTGAGCGATGCCAC 3240
Db 3181 TATCAGCAGCGCTTCTGGGGACCATCTGTGATGACGCTGGGACCTGAGCGATGCCAC 3240
Qy 3241 GTGGTGTCTAAAAGCTGGGCTGTGGAGTGGCCCTCAATGCCAGGCTCTGCTCAGCTTT 3300
Db 3241 GTGGTGTCTAAAAGCTGGGCTGTGGAGTGGCCCTCAATGCCAGGCTCTGCTCAGCTTT 3300
Qy 3301 GGGGAGGGGTTCAGGCGCCATCTGGCTGGATGACCTGAACTGCACAGAAACGGAGTCCAC 3360
Db 3301 GGGGAGGGGTTCAGGCGCCATCTGGCTGGATGACCTGAACTGCACAGAAACGGAGTCCAC 3360
Qy 3361 TTGTGGCAGTCCCTTCCCGGGTGGGGGACGACGACTGCAGGCAACAGGAGGACCCA 3420
Db 3361 TTGTGGCAGTCCCTTCCCGGGTGGGGGACGACGACTGCAGGCAACAGGAGGACCCA 3420
Qy 3421 GGGGTCTGCTCAGAAATTCACAGCCTTGAGGCTCTACAGTGAACACTGAAACAGAGAGC 3480
Db 3421 GGGGTCTGCTCAGAAATTCACAGCCTTGAGGCTCTACAGTGAACACTGAAACAGAGAGC 3480

Qy 3481 TGTCTGGAGATTGGAAGTCTTCTATTAACGGAGCTTGGGACGCTGGCAGGAGGAAC 3540
Db 3481 TGTCTGGAGATTGGAAGTCTTCTATTAACGGAGCTTGGGACGCTGGCAGGAGGAAC 3540
Qy 3541 ATCACCACAGCCATAGCAGGCAATTTGTGTGACAGGAGCTGGGCTGTGGGAGAAATGGAGTT 3600
Db 3541 ATCACCACAGCCATAGCAGGCAATTTGTGTGACAGGAGCTGGGCTGTGGGAGAAATGGAGTT 3600
Qy 3601 CTCAGCCTCGCCCTTTATCTAAAGACAGGCTCTGGTTTCAATGTGGGTGATGACATTCAG 3660
Db 3601 CTCAGCCTCGCCCTTTATCTAAAGACAGGCTCTGGTTTCAATGTGGGTGATGACATTCAG 3660
Qy 3661 TGTCTTAAACGCATATCTCCATATGTCAGTGCCTGTCTGCCCATGGAGCGAAGAATC 3720
Db 3661 TGTCTTAAACGCATATCTCCATATGTCAGTGCCTGTCTGCCCATGGAGCGAAGAATC 3720
Qy 3721 TCCAGCCACAGCAGAGACCTGGATCATCATGTGAAGATAGAAATAGAGTGGGTGGAGGA 3780
Db 3721 TCCAGCCACAGCAGAGACCTGGATCATCATGTGAAGATAGAAATAGAGTGGGTGGAGGA 3780
Qy 3781 GACACCGAGTCTCTGGGAGAGTGGAGATCTGGCACGAGGCTCTGGGGCACAGTGTGT 3840
Db 3781 GACACCGAGTCTCTGGGAGAGTGGAGATCTGGCACGAGGCTCTGGGGCACAGTGTGT 3840
Qy 3841 CATGACTCTCTGGGACCTGGCCGAGCGGAAGTGTGTGTCTACAGAGCTGGGCTGTGGCTCT 3900
Db 3841 CATGACTCTCTGGGACCTGGCCGAGCGGAAGTGTGTGTCTACAGAGCTGGGCTGTGGCTCT 3900
Qy 3901 GCTCTGGCTGGCTCTGAGGAGCGCTTCGTTTGGCCAGGAACTGGAACCATCTGTTGGAT 3960
Db 3901 GCTCTGGCTGGCTCTGAGGAGCGCTTCGTTTGGCCAGGAACTGGAACCATCTGTTGGAT 3960
Qy 3961 GACATCGGCTGCAAGAGAAATGAGTCAATTTCTATGGGACTGTACGCCAACCCCTGGGGA 4020
Db 3961 GACATCGGCTGCAAGAGAAATGAGTCAATTTCTATGGGACTGTACGCCAACCCCTGGGGA 4020
Qy 4021 CAGAGTGAATGCTGAGGACACAGGAAGATGCTGGCGTGGAGTGTCTGGACAGTGCCTGAAA 4080
Db 4021 CAGAGTGAATGCTGAGGACACAGGAAGATGCTGGCGTGGAGTGTCTGGACAGTGCCTGAAA 4080
Qy 4081 TCAGTGAATGCTGCTCAGGTCATTTAGCACTTATTTATCCAGTATCTTTGGGCTCCTT 4140
Db 4081 TCAGTGAATGCTGCTCAGGTCATTTAGCACTTATTTATCCAGTATCTTTGGGCTCCTT 4140
Qy 4141 CTCTGGTTCGTTTATTTCTATTTCTCAGTGTGGCTGCGAGTTTCAGAAACAAACATCTG 4200
Db 4141 CTCTGGTTCGTTTATTTCTATTTCTCAGTGTGGCTGCGAGTTTCAGAAACAAACATCTG 4200
Qy 4201 CCCCTCAGAGTTTCAACAGAGAGGAGGTTCTCTCGAGGAGAAATTTATTCATGAGATG 4260
Db 4201 CCCCTCAGAGTTTCAACAGAGAGGAGGTTCTCTCGAGGAGAAATTTATTCATGAGATG 4260
Qy 4261 GAGAGCTGCCCTCAAGAGAGAGGACCCACATGGGACAGAACCTCAGATGACACCCCAAC 4320
Db 4261 GAGAGCTGCCCTCAAGAGAGAGGACCCACATGGGACAGAACCTCAGATGACACCCCAAC 4320
Qy 4321 CATGCTCTGAAGATGCTAGGACACATCGCTGTGGGAGTTCCTCTCCCTCTCAAGCC 4380
Db 4321 CATGCTCTGAAGATGCTAGGACACATCGCTGTGGGAGTTCCTCTCCCTCTCAAGCC 4380
Qy 4381 ACAAAATGACTTTAGCACTTCCAGGCTCACCAGATCAACCTCTAAATATCTTTGAAGAG 4440
Db 4381 ACAAAATGACTTTAGCACTTCCAGGCTCACCAGATCAACCTCTAAATATCTTTGAAGAG 4440
Qy 4441 ACACAACTTTTAAATGAATGAAGAGAGTCAAGTTCGCCCTATGGGAAACCTGTGCCAAA 4500
Db 4441 ACACAACTTTTAAATGAATGAAGAGAGTCAAGTTCGCCCTATGGGAAACCTGTGCCAAA 4500
Qy 4501 TAACTTTCTTGAACAATAGGAGAACAGTAAATTTGATTAAGAGCTGGTGAATAAATAAT 4560
Db 4501 TAACTTTCTTGAACAATAGGAGAACAGTAAATTTGATTAAGAGCTGGTGAATAAATAAT 4560

Db	1201	ACAAATATGTCACCAGAACTGGGAAGAAATGAACAAGCCCTTGTGGTTTGTAAACGACGCTAGGA	1261
Qy	1261	TGTCGGTTACAGCTCTTTGGCAGTCTGCTGCTAAACCTAGTAATGAAGCTAGAGACATT	1320
Db	1261	TGTCGGTTACAGCTCTTTGGCAGTCTGCTGCTAAACCTAGTAATGAAGCTAGAGACATT	1320
Qy	1321	TGGATAAACAGCATATCTTGCACCTGGGAATGAGTCAGCTCTCTGGGACTGCACATATGAT	1380
Db	1321	TGGATAAACAGCATATCTTGCACCTGGGAATGAGTCAGCTCTCTGGGACTGCACATATGAT	1380
Qy	1381	GGAAAGCAAGCGACATATCTTCCGAGATCAGATGCTGGAGTAATTTGTCTTGATAAG	1440
Db	1381	GGAAAGCAAGCGACATATCTTCCGAGATCAGATGCTGGAGTAATTTGTCTTGATAAG	1440
Qy	1441	GCAGATCTGGACCTAAAGGCTTGTGGGGGCTCATAGCCCTGTTATGGGAGATGGAGGTG	1500
Db	1441	GCAGATCTGGACCTAAAGGCTTGTGGGGGCTCATAGCCCTGTTATGGGAGATGGAGGTG	1500
Qy	1501	AAATACCAAGGAGTGGGGAGCTGTGTCTCATGACAGATGGAGCAACAAGGAATGCAAGCT	1560
Db	1501	AAATACCAAGGAGTGGGGAGCTGTGTCTCATGACAGATGGAGCAACAAGGAATGCAAGCT	1560
Qy	1561	GTTGTGTGTAACAATTTGGGATCTGGAAGCCATGTCATGCTCTTGGTATGACCTATTTT	1620
Db	1561	GTTGTGTGTAACAATTTGGGATCTGGAAGCCATGTCATGCTCTTGGTATGACCTATTTT	1620
Qy	1621	AAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTCTTGCAATTTGAAATGAGTCAAAAT	1680
Db	1621	AAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTCTTGCAATTTGAAATGAGTCAAAAT	1680
Qy	1681	ATCTGGGACTGTGAACACAGTGGATGGGAAAGCATAAATCTGTACACAGAGAGGATGTG	1740
Db	1681	ATCTGGGACTGTGAACACAGTGGATGGGAAAGCATAAATCTGTACACAGAGAGGATGTG	1740
Qy	1741	ATTGTAACCTGCTCAGGTGATGCAACATGGGGCCCTGAGGCTGGTGGGGCGCAGCAACCGC	1800
Db	1741	ATTGTAACCTGCTCAGGTGATGCAACATGGGGCCCTGAGGCTGGTGGGGCGCAGCAACCGC	1800
Qy	1801	TGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTGGGGCACAGTGTGTGATGACGGC	1860
Db	1801	TGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTGGGGCACAGTGTGTGATGACGGC	1860
Qy	1861	TGGAACAGTAAAGCTCAGCTGTGGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATTT	1920
Db	1861	TGGAACAGTAAAGCTCAGCTGTGGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATTT	1920
Qy	1921	GGCATGGGCTCGGGAACGCTTCTACAGATATGAAAAAATTTGGCTCGATGATGTTTCC	1980
Db	1921	GGCATGGGCTCGGGAACGCTTCTACAGATATGAAAAAATTTGGCTCGATGATGTTTCC	1980
Qy	1981	TGTGATGGAGATCAGTCAGATCTCTGGTCTATCGAGAAACAGTGGGTGGGAAATTAATGAC	2040
Db	1981	TGTGATGGAGATCAGTCAGATCTCTGGTCTATCGAGAAACAGTGGGTGGGAAATTAATGAC	2040
Qy	2041	TGCAGTCCACAGTGAAGATGTGGAGTGAATCTGTTCTGATGATCGGATATGAGGCTGAGG	2100
Db	2041	TGCAGTCCACAGTGAAGATGTGGAGTGAATCTGTTCTGATGATCGGATATGAGGCTGAGG	2100
Qy	2101	CTTGTGGGTGGAGCAGCAGGTGCTCGAAAGATTGAGGTGAATGTCACAGGCTGCCGTG	2160
Db	2101	CTTGTGGGTGGAGCAGCAGGTGCTCGAAAGATTGAGGTGAATGTCACAGGCTGCCGTG	2160
Qy	2161	GGAAATCTGTGTCTAATGGCTGGGAATGAACATTTGCTGAAGTTGTTTTCAGGCAACTT	2220
Db	2161	GGAAATCTGTGTCTAATGGCTGGGAATGAACATTTGCTGAAGTTGTTTTCAGGCAACTT	2220
Qy	2221	GAATGTGGGTCTGCAATCAGGCTCTCCAGAGAGCCTCATTTCCAGAAAGAACATTAAC	2280
Db	2221	GAATGTGGGTCTGCAATCAGGCTCTCCAGAGAGCCTCATTTCCAGAAAGAACATTAAC	2280
Qy	2281	ATCTTAATGTGAAATTTCTGGCTGCACCTGGAGGGAGGCTCTCTCTGGGATTTGATACGA	2340
Db	2281	ATCTTAATGTGAAATTTCTGGCTGCACCTGGAGGGAGGCTCTCTCTGGGATTTGATACGA	2340

QY	2341	TGGAGTGGAAACAGACTCGGTGTCATTTAAATATATGGAAGCAAGTTTGATCTCTCTCAGCC	2400
DB	2341	TGGGAGTGGAAACAGACTCGGTGTCATTTAAATATATGGAAGCAAGTTTGATCTCTCTCAGCC	2400
QY	2401	CACAGCAGCCAGCGTCTGGAGCTGATATGCCCTGCTCTGAGCTGTGGAAGTGAAA	2460
DB	2401	CACAGCAGCCAGCGTCTGGAGCTGATATGCCCTGCTCTGAGCTGTGGAAGTGAAA	2460
QY	2461	CATGCAGACACATGGCGCTCTGTCTGTGATCTTGATTTCTCTTCTTCATGCTGCCAATGTG	2520
DB	2461	CATGCAGACACATGGCGCTCTGTCTGTGATCTTGATTTCTCTTCTTCATGCTGCCAATGTG	2520
QY	2521	CTGTGCAGAGAAATTAATTTGTGGAGATGCCATATCTCTTCTGTGGGAGATCACTTTGGA	2580
DB	2521	CTGTGCAGAGAAATTAATTTGTGGAGATGCCATATCTCTTCTGTGGGAGATCACTTTGGA	2580
QY	2581	AAAGGGAATGGTCTAACTTTGGGCCGAAAAAGTTCCAGTGTGAAGGGAGTGAACACTCACCTT	2640
DB	2581	AAAGGGAATGGTCTAACTTTGGGCCGAAAAAGTTCCAGTGTGAAGGGAGTGAACACTCACCTT	2640
QY	2641	GCATTTATGCCCATTTGTTTCAACATCCGGAGACACTTTGTATCCACAGCAGAGAAGTTGGA	2700
DB	2641	GCATTTATGCCCATTTGTTTCAACATCCGGAGACACTTTGTATCCACAGCAGAGAAGTTGGA	2700
QY	2701	GTTTGTCTGTTCCCGATATACAGATGTCGCACTTGTGAATGGCAAAATCCCACTGTGACGGG	2760
DB	2701	GTTTGTCTGTTCCCGATATACAGATGTCGCACTTGTGAATGGCAAAATCCCACTGTGACGGG	2760
QY	2761	CAAGTGGAGATCAACGTGTTGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCA	2820
DB	2761	CAAGTGGAGATCAACGTGTTGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCA	2820
QY	2821	GAAGATGCCCGTCTATGTCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCCACAGGA	2880
DB	2821	GAAGATGCCCGTCTATGTCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACCCACAGGA	2880
QY	2881	CGAAATATATTGGAAAAAGTGTTCGTGTGTGGGGACACAGGTTTCATTCCTTTAGGG	2940
DB	2881	CGAAATATATTGGAAAAAGTGTTCGTGTGTGGGGACACAGGTTTCATTCCTTTAGGG	2940
QY	2941	AATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCCTTGGAGCACCCTCCTGTATCCAT	3000
DB	2941	AATGAGTCACTTCTGGATAACTGTCAAATGACAGTTCCTTGGAGCACCCTCCTGTATCCAT	3000
QY	3001	GGAAATACATCTCTGTGATCTGCACAGGAAGCCTGACCCAGGCACTGTTTCCATGCCTC	3060
DB	3001	GGAAATACATCTCTGTGATCTGCACAGGAAGCCTGACCCAGGCACTGTTTCCATGCCTC	3060
QY	3061	GCAAATGTATCTGACCCATATTTGCTCGAGTTCCAGAGGGCACTGTTTGATCTGCTTA	3120
DB	3061	GCAAATGTATCTGACCCATATTTGCTCGAGTTCCAGAGGGCACTGTTTGATCTGCTTA	3120
QY	3121	GAGGACAAACGGCTCCGCCCTAGTGGATGGGACAGCGCTGTCCGGGAGAGTAGAGATC	3180
DB	3121	GAGGACAAACGGCTCCGCCCTAGTGGATGGGACAGCGCTGTCCGGGAGAGTAGAGATC	3180
QY	3181	TATCAGCAGCGTCTTGGGGACCACTCTGTGTATGACGGTGGGACCTTGACGATGCCAC	3240
DB	3181	TATCAGCAGCGTCTTGGGGACCACTCTGTGTATGACGGTGGGACCTTGACGATGCCAC	3240
QY	3241	GTGGTGTGTCAAAGCTGGGCTGTGAGTGGCCCTTCAATGCCACGGTCTCTGCTCACTTT	3300
DB	3241	GTGGTGTGTCAAAGCTGGGCTGTGAGTGGCCCTTCAATGCCACGGTCTCTGCTCACTTT	3300
QY	3301	GGGGAGGGCTCAGGGGCCATCTGGCTGGATGACCTTGAACCTGCACAGGAACGAGTCCCAC	3360
DB	3301	GGGGAGGGCTCAGGGGCCATCTGGCTGGATGACCTTGAACCTGCACAGGAACGAGTCCCAC	3360
QY	3361	TTGTGGCAGTGCCCTTCCCGCGGCTGGGGCAGCAGCTGCAAGGCACAAAGGAGGACGA	3420
DB	3361	TTGTGGCAGTGCCCTTCCCGCGGCTGGGGCAGCAGCTGCAAGGCACAAAGGAGGACGA	3420

Qy 508 AACAACTCCCTGTTTCAGGAGAGTGGAGGTGAAATTTCCAAGAAAGTGGGGACTATATGT 567
Db 481 AACAACTCCCTGTTTCAGGAGAGTGGAGGTGAAATTTCCAAGAAAGTGGGGACTATATGT 540
Qy 568 GATGATGGGTGGAACATTGAATACATGCTGCCGTGGTGTGCGAGGCAACATGAGATGCCATCT 627
Db 541 GATGATGGGTGGAACATTGAATACATGCTGCCGTGGTGTGCGAGGCAACATGAGATGCCATCT 600
Qy 628 TCCTTTATTTCTCTGAGTGTGTAATAGCCCTGCTGTATGCGCCCATTTGGCTGGAT 687
Db 601 TCCTTTATTTCTCTGAGTGTGTAATAGCCCTGCTGTATGCGCCCATTTGGCTGGAT 660
Qy 688 GACATTTATTTCCAGAGGGAATGAGTGGCACTCTGGAATTCGACACATCTGGTGGATGGGA 747
Db 661 GACATTTATGCGAGGGAATGAGTGGCACTCTGGAATTCGACACATCTGGTGGATGGGA 720
Qy 748 AATCATGCTGCACTACAATGAGGATGTCACATTAATCTGTTATGATAGTAGTATCTT 807
Db 721 AATCATGCTGCACTACAATGAGGATGTCACATTAATCTGTTATGATAGTAGTATCTT 780
Qy 808 GAACATAAGGCTTTAGTGGAACTAACCGCTGTATGCGGAGAGTAGAGCTGAAATCCAA 867
Db 781 GAACATAAGGCTTTAGTGGAACTAACCGCTGTATGCGGAGAGTAGAGCTGAAATCCAA 840
Qy 868 GGAAGTGGGGACCGTATGCCACATAAGTGGAAACAATGCTGCAGCTGATGCTGATGC 927
Db 841 GGAAGTGGGGACCGTATGCCACCATAGTGGAAACAATGCTGCAGCTGATGCTGATGC 900
Qy 928 AAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGCTGCTGCTCAATTTGCACTCAGG 987
Db 901 AAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGCTGCTGCTCAATTTGCACTCAGG 960
Qy 988 TCTCATCTGTATGCTGTGATGCTGCTCTGCTCCGATGATGATCTTTCTTTGGGAC 1047
Db 961 TCTCATCTGTATGCTGTGATGCTGCTCTGCTCCGATGATGATCTTTCTTTGGGAC 1020
Qy 1048 TGCAGACATTTCCGGAACCGTCAATTTTGACTGCTCTCATCAAAACGATGCTGCTGATC 1107
Db 1021 TGCAGACATTTCCGGAACCGTCAATTTTGACTGCTCTCATCAAAACGATGCTGCTGATC 1080
Qy 1108 TGCTCAGATGAGCAGATTTGGAACTCGGACTAGCAGATGAGAAATACAATTTGTCAGG 1167
Db 1081 TGCTCAGATGAGCAGATTTGGAACTCGGACTAGCAGATGAGAAATACAATTTGTCAGG 1140
Qy 1168 AGACTAGAGTGGAATTCATGACAGTGTGGACATATGTGACCAAGCACTGCAAGAT 1227
Db 1141 AGACTAGAGTGGAATTCATGACAGTGTGGACATATGTGACCAAGCACTGCAAGAT 1200
Qy 1228 GAACAAGCCCTTGTGGTTTGTAAAGCAGCTAGGATGCTCGTTTCAGCGCTTTGGCAGTCTG 1287
Db 1201 GAACAAGCCCTTGTGGTTTGTAAAGCAGCTAGGATGCTCGTTTCAGCGCTTTGGCAGTCTG 1260
Qy 1288 CGTCTAAACCTAGTAATGAAGCTAGAGACATTTGGATTAACAGCATACTTTGCACTGGG 1347
Db 1261 CGTCTAAACCTAGTAATGAAGCTAGAGACATTTGGATTAACAGCATACTTTGCACTGGG 1320
Qy 1348 AATCAGTCAGCTCTCTGGACTGCACATATGATGGAAGCAACAGCAACATGCTTCGA 1407
Db 1321 AATCAGTCAGCTCTCTGGACTGCACATATGATGGAAGCAACAGCAACATGCTTCGA 1380
Qy 1408 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTTAAGCCTTTGTCGGG 1467
Db 1381 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTTAAGCCTTTGTCGGG 1440
Qy 1468 GCTCATAGCCCTGTTATGGGAGATTTGGAGTGAAATACCAAGAGAGTGGGGACTGTG 1527
Db 1441 GCTCATAGCCCTGTTATGGGAGATTTGGAGTGAAATACCAAGAGAGTGGGGACTGTG 1500
Qy 1528 TGTATGACAGATGGACACAGGAATCCAGCTGTTGCTGTAACAAATTTGGGATGTGGA 1587
Db 1501 TGTATGACAGATGGACACAGGAATCCAGCTGTTGCTGTAACAAATTTGGGATGTGGA 1560

Qy 1588 AAGCCTATGATGTGTTTGGTATGACCTATTTTAAAGAACATCAGGACCTATTTTGGCTG 1647
Db 1561 AAGCCTATGATGTGTTTGGTATGACCTATTTTAAAGAACATCAGGACCTATTTTGGCTG 1620
Qy 1648 GATGACGTTTCTTCATTTGGAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGG 1707
Db 1621 GATGACGTTTCTTCATTTGGAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGG 1680
Qy 1708 GGAAGCATAAATCTGTACACAGAGAGGATGTGATTTAACTGCTCAGGTGATGCAACA 1767
Db 1681 GGAAGCATAAATCTGTACACAGAGAGGATGTGATTTAACTGCTCAGGTGATGCAACA 1740
Qy 1768 TGGGCTCTGAGCTGTGGGCGGAGCAACCGCTGCTCGGAAGACTCGAGGTGACTTTT 1827
Db 1741 TGGGCTCTGAGCTGTGGGCGGAGCAACCGCTGCTCGGAAGACTCGAGGTGACTTTT 1800
Qy 1828 CAAGACGCTGGGCGACAGTGTGTATGACGCTGGAACAGTAAAGCTGCAGCTGTGGTG 1887
Db 1801 CAAGACGCTGGGCGACAGTGTGTATGACGCTGGAACAGTAAAGCTGCAGCTGTGGTG 1860
Qy 1888 TGTAGCCAGCTGGAGCTGCCCATCTTCTATCTATGTCATGCTGGGCTGCGGAAACGCTTCTACA 1947
Db 1861 TGTAGCCAGCTGGAGCTGCCCATCTTCTATCATTTGGCATGGGTCTGGGAAACGCTTCTACA 1920
Qy 1948 GGATATGGAATAATTTGGCTCGATGATGTTTCTCTGTGATGAGATGAGTCAAGTCTCTGG 2007
Db 1921 GGATATGGAATAATTTGGCTCGATGATGTTTCTCTGTGATGAGATGAGTCAAGTCTCTGG 1980
Qy 2008 TCATGACGGAACAGTGGGTGGGGAATAATGACTGCTCAGTCACAGTGAAGATGTTGGAGTG 2067
Db 1981 TCATGACGGAACAGTGGGTGGGGAATAATGACTGCTCAGTCACAGTGAAGATGTTGGAGTG 2040
Qy 2068 ATCTGTTCTGATGATCGATCGATATGAGCTGAGGCTTTGGGTGGAAGCAGCAGTGTGCT 2127
Db 2041 ATCTGTTCTGATGATCGATCGATATGAGCTGAGGCTTTGGGTGGAAGCAGCAGTGTGCT 2100
Qy 2128 GGAAGAAGTTGAGGTGAATGTCAGGGTCCGCTGGGAATTCGTGTGCTTAATGGCTGGGGA 2187
Db 2101 GGAAGAAGTTGAGGTGAATGTCAGGGTCCGCTGGGAATTCGTGTGCTTAATGGCTGGGGA 2160
Qy 2188 ATGAACATTTGCTGAAGTTGTTTTCAGGCACTTGAATGTGGTCTGCAATCAGGCTCTCC 2247
Db 2161 ATGAACATTTGCTGAAGTTGTTTTCAGGCACTTGAATGTGGTCTGCAATCAGGCTCTCC 2220
Qy 2248 AGAGAGCCTCATTTTCACAGAAGAACATTTACACATCTTAAATGCTGAAATCTTGGCTGCACT 2307
Db 2221 AGAGAGCCTCATTTTCACAGAAGAACATTTACACATCTTAAATGCTGAAATCTTGGCTGCACT 2280
Qy 2308 GGAGGGAAAGCCTCTCTCTGGGATTTGATACGATGGGAGTGGAAACAGACTGCGTGTCTAT 2367
Db 2281 GGAGGGAAAGCCTCTCTCTGGGATTTGATACGATGGGAGTGGAAACAGACTGCGTGTCTAT 2340
Qy 2368 TTAATATGGAAGCAAGTTTGTATCTGCTGAGTTCAGCCACAGCAGCCAGGCTGGTGGAGCT 2427
Db 2341 TTAATATGGAAGCAAGTTTGTATCTGCTGAGTTCAGCCACAGCAGCCAGGCTGGTGGAGCT 2400
Qy 2428 GATATGCCCTGCTCTGAGCTGTTTGAAGTGAACATGACAGACATGCGCTCTCTGCTGT 2487
Db 2401 GATATGCCCTGCTCTGAGCTGTTTGAAGTGAACATGACAGACATGCGCTCTCTGCTGT 2460
Qy 2488 GATCTGATTTCTCTCTCATGCTGCCAATGTGCTGTCAGAGAAATTAATTTGTTGGAGAT 2547
Db 2461 GATCTGATTTCTCTCTCATGCTGCCAATGTGCTGTCAGAGAAATTAATTTGTTGGAGAT 2520
Qy 2548 GCCATATCTCTCTCTGTTGGGAGATCACTTTGGAAGGGAATGCTTAACTTTGGGCCGAA 2607
Db 2521 GCCATATCTCTCTCTGTTGGGAGATCACTTTGGAAGGGAATGCTTAACTTTGGGCCGAA 2580
Qy 2608 RAGTTCAGTGTGAAGGAGTGAACCTCAGCTTGCATTTATGCCCATTTGTTCAACATCCG 2667
Db 2581 RAGTTCAGTGTGAAGGAGTGAACCTCAGCTTGCATTTATGCCCATTTGTTCAACATCCG 2640
Qy 2668 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGCTGTTCCCGATATACAGATGTC 2727

Db	2641		GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGCTGTCTCCCGAATATACAGATGTC	2700
Qy	2728	CGACTTGTGAATGGCAAAATCCCAAGTGTACGGCAAGTGGAGATCAACGTGCTTGACAC	2787	
Db	2701	CGACTTGTGAATGGCAAAATCCCAAGTGTACGGCAAGTGGAGATCAACGTGCTTGACAC	2760	
Qy	2788	TGGGGCTCACTGTGTACACCCACCTGGGACCCAGAAGATGCCGTGTTCTATGCACACAG	2847	
Db	2761	TGGGGCTCACTGTGTACACCCACCTGGGACCCAGAGATGCCGTGTTCTATGCACACAG	2820	
Qy	2848	CTCAGCTGTGGAGTGCTCTCTCAACCCACAGAGGAAATATATTGGAGAAAGAGTGT	2907	
Db	2821	CTCAGCTGTGGAGTGCTCTCTCAACCCACAGAGGAAATATATTGGAGAAAGAGTGT	2880	
Qy	2908	CGTGTGTGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACCTTCGGATAACTGTCAA	2967	
Db	2881	CGTGTGTGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACCTTCGGATAACTGTCAA	2940	
Qy	2968	ATCACAGTTCCTTGGAGCACTCCCTGTATCCATGGAATACTGTCGTGATCTGCACA	3027	
Db	2941	ATCACAGTTCCTTGGAGCACTCCCTGTATCCATGGAATACTGTCGTGATCTGCACA	3000	
Qy	3028	GGAAAGCCTGACCCAGCCACTGTTTCCATGCCTTCGCAATGATCTGACCCATATTTGTCT	3087	
Db	3001	GGAAAGCCTGACCCAGCCACTGTTTCCATGCCTTCGCAATGATCTGACCCATATTTGTCT	3060	
Qy	3088	GCAGTTCACAGAGGCACTGCTTTGATCTGCTTAGAGGACAACGGCTCCGCCCTAGTGGAT	3147	
Db	3061	GCAGTTCACAGAGGCACTGCTTTGATCTGCTTAGAGGACAACGGCTCCGCCCTAGTGGAT	3120	
Qy	3148	GGGACAGCGCTGTGCCGGGAGTAGAGATCTATCACGACGGCTCTCTGGGCAACCATC	3207	
Db	3121	GGGACAGCGCTGTGCCGGGAGTAGAGATCTATCACGACGGCTCTCTGGGCAACCATC	3180	
Qy	3208	TGTGATGACGGCTGGGACCTGAGCGATGCCACGTGTGTGTCAAAAGCTGGGCTGTGGA	3267	
Db	3181	TGTGATGACGGCTGGGACCTGAGCGATGCCACGTGTGTGTCAAAAGCTGGGCTGTGGA	3240	
Qy	3268	GTGGCTTCAATGCCACGCTCTGCTCACTTTGGGAGGGTCAGGGCCCATCTGGCTG	3327	
Db	3241	GTGGCTTCAATGCCACGCTCTGCTCACTTTGGGAGGGTCAGGGCCCATCTGGCTG	3300	
Qy	3328	GATGACCTGAACCTGCACAGAAACGGAGTCCCATTGTGGCAGTGCCTTCCCGCGCTGG	3387	
Db	3301	GATGACCTGAACCTGCACAGAAACGGAGTCCCATTGTGGCAGTGCCTTCCCGCGCTGG	3360	
Qy	3388	GGCAGCAGCTGTCAGGCACAAGGAGGACGAGGGGTATCTGCTCAGAAATTCACAGCC	3447	
Db	3361	GGCAGCAGCTGTCAGGCACAAGGAGGACGAGGGGTATCTGCTCAGAAATTCACAGCC	3420	
Qy	3448	TTGAGGCTCTACAGTCAAACTCAAAACAGAGAGCTGTGCTGGAGCATTTGGAAGTCTTCAT	3507	
Db	3421	TTGAGGCTCTACAGTCAAACTCAAAACAGAGAGCTGTGCTGGAGCATTTGGAAGTCTTCAT	3480	
Qy	3508	AACGGACCTTGGGCAACGCTCGGCAGGAGGAACATCACACAGCCATAGCAGGCATTGTG	3567	
Db	3481	AACGGACCTTGGGCAACGCTCGGCAGGAGGAACATCACACAGCCATAGCAGGCATTGTG	3540	
Qy	3568	TGCAGGCACTGGGCTGTGGGAGAAATGGAGTTGTACGCTTCGCCCTTTATCTAAGACA	3627	
Db	3541	TGCAGGCACTGGGCTGTGGGAGAAATGGAGTTGTACGCTTCGCCCTTTATCTAAGACA	3600	
Qy	3628	GGCTCTGTTTCAATGTGGGTGATGACATTCAGTGTCTCTAAACGCATATCTCCATATGG	3687	
Db	3601	GGCTCTGTTTCAATGTGGGTGATGACATTCAGTGTCTCTAAACGCATATCTCCATATGG	3660	
Qy	3688	CAGTGCTGTCTGCCCATGGGAGGAGAAATCTCCAGCCACAGAGACACCTGGATC	3747	
Db	3661	CAGTGCTGTCTGCCCATGGGAGGAGAAATCTCCAGCCACAGAGACACCTGGATC	3720	
Qy	3748	ACATGTGAAGATAGAAATAGTGTGGAGAGACACCGAGTGCTCTGGGAGAGTGAG	3807	

Db	3721	ACATGTGAGATAGAAATAGACTGCGTGGAGAGACACCGAGTGCTCTGGGAGAGTGGAG	3780
Qy	3808	ATCTGGCACGACAGCTTCCTGGGCGACAGTGTGTATGACTCCTGGGACCTGGCCGAGGCG	3867
Db	3781	ATCTGGCACGACAGCTTCCTGGGCGACAGTGTGTATGACTCCTGGGACCTGGCCGAGGCG	3840
Qy	3868	GAAGTGGTGTTCACGAGCTGGGCTGTGGCTCTGCTCTGGCTGGCCCTGAGGAGCGCTTCG	3927
Db	3841	GAAGTGGTGTTCACGAGCTGGGCTGTGGCTCTGCTCTGGCTGGCCCTGAGGAGCGCTTCG	3900
Qy	3928	TTTGGCCAGGGAACCTGGAACCATCTGGTGTGGATGACATGCGGTGCARAGGAATGACTCA	3987
Db	3901	TTTGGCCAGGGAACCTGGAACCATCTGGTGTGGATGACATGCGGTGCARAGGAATGACTCA	3960
Qy	3988	TTTCTATGGGACTGTCCAGCCAAACCCCTGGGCGACAGAGTGACTGTGGACACAGAAGAT	4047
Db	3961	TTTCTATGGGACTGTCCAGCCAAACCCCTGGGCGACAGAGTGACTGTGGACACAGAAGAT	4020
Qy	4048	GCTGGGCTGAGTGTCTCTGGACAGTCGCTGGAATCACTGAATGCGCTCTCAGTCTATTTA	4107
Db	4021	GCTGGGCTGAGTGTCTCTGGACAGTCGCTGGAATCACTGAATGCGCTCTCAGTCTATTTA	4080
Qy	4108	GCATCTATTTTATCCAGTATCTTTGGGCTCCCTTCCTGGTTCGTCTTATTCTTATTTCTC	4167
Db	4081	GCATCTATTTTATCCAGTATCTTTGGGCTCCCTTCCTGGTTCGTCTTATTCTTATTTCTC	4140
Qy	4168	ACGTGTGCGGAGTTCCAGAAACAAAACATCTGCCCTCAGAGTTTCAACCAAGAGGAGG	4227
Db	4141	ACGTGTGCGGAGTTCCAGAAACAAAACATCTGCCCTCAGAGTTTCAACCAAGAGGAGG	4200
Qy	4228	GGTTCTCTCAGAGAGATTTATTTCATGAGATGGAGACCTGCCTCAAGAGAGAGACCCA	4287
Db	4201	GGTTCTCTCAGAGAGATTTATTTCATGAGATGGAGACCTGCCTCAAGAGAGAGACCCA	4260
Qy	4288	CATGGGCAAGAACCCTCAGATGACACCCCAACCATGGTGTGGAAGATGCTTAGCGACACA	4347
Db	4261	CATGGGCAAGAACCCTCAGATGACACCCCAACCATGGTGTGGAAGATGCTTAGCGACACA	4320
Qy	4348	TCGCTGTGGGAGTTCTTCTGCCTCTGAAGCCACAAA	4386
Db	4321	TCGCTGTGGGAGTTCTTCTGCCTGAAGCCACAAA	4359

RESULT 4

US-09-759-130B-380

Sequence 380, Application US/09759130B

Publication No. US20030022279A1

GENERAL INFORMATION:

APPLICANT: Millennium Pharmaceuticals, Inc.

APPLICANT: McCarthy, Sean A

APPLICANT: Fraser, Christopher C

APPLICANT: Sharp, John D

APPLICANT: Barnes, Thomas S

APPLICANT: Kirst, Susan J

APPLICANT: Mackay, Charles R

APPLICANT: Myers, Paul S

APPLICANT: Leiby, Kevin R

APPLICANT: Wrighton, Nicolas

APPLICANT: Goodearl, Andrew

APPLICANT: Holtzman, Douglas A

TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING

TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE,

TITLE OF INVENTION: USES.

FILE REFERENCE: MPI00-5350MIM

CURRENT APPLICATION NUMBER: US/09/759,130B

CURRENT FILING DATE: 2002-09-16

PRIOR APPLICATION NUMBER: US 09/479,249

PRIOR FILING DATE: 2000-01-07

PRIOR APPLICATION NUMBER: US 09/559,497

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: US 09/578,063

PRIOR FILING DATE: 2000-05-24

PRIOR APPLICATION NUMBER: US 09/333,159

|||||
Db 1801 CRAAGGCGTGGGGCACAGTGTGTGATGACGGCTGGAAACAGTAAAGCTGCAGCTGTGGTG 1860
QY 1888 TGTAGCAGCTGACGACGACCTCTTCTATCATTTGGCATGGGTCTGGAAACGCTTCTACA 1947
Db 1861 TGTAGCAGCTGACGACGACCTCTTCTATCATTTGGCATGGGTCTGGAAACGCTTCTACA 1920
QY 1948 GGATATGGAAAAATTTGGCTCGCATGATGTTTCTGTGTGATGGAGATGAGTCAGATCTCTGG 2007
Db 1921 GGATATGGAAAAATTTGGCTCGCATGATGTTTCTGTGTGATGGAGATGAGTCAGATCTCTGG 1980
QY 2008 TCATGACGAGAACAGTGGGTGGGAAATATGACTGCAGTTCACAGTCAAGATGTTGAGTG 2067
Db 1981 TCATGACGAGAACAGTGGGTGGGAAATATGACTGCAGTTCACAGTCAAGATGTTGAGTG 2040
QY 2068 ATCTGTCTCTGATCATCGCATGATGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCT 2127
Db 2041 ATCTGTCTCTGATCATCGCATGATGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCT 2100
QY 2128 GGAAGAGTTGAGGTGAATGTCCAGGCTGCCGTGGGAATTTCTGTGCTTAATGGCTGGGA 2187
Db 2101 GGAAGAGTTGAGGTGAATGTCCAGGCTGCCGTGGGAATTTCTGTGCTTAATGGCTGGGA 2160
QY 2188 ATGAACATTGCTCAAGTTGTTTCAGGCAACTTGAATGTGGGTCTGCAATCAGGCTCTCC 2247
Db 2161 ATGAACATTGCTCAAGTTGTTTCAGGCAACTTGAATGTGGGTCTGCAATCAGGCTCTCC 2220
QY 2248 AGAGAGCTCATTTACAGAAAGAACATTACACATCTTAATGTCGAATTTCTGGCTGCAT 2307
Db 2221 AGAGAGCTCATTTACAGAAAGAACATTACACATCTTAATGTCGAATTTCTGGCTGCAT 2280
QY 2308 GGAGGGAAGCCCTCTCTCTGGGATTTATACGATGGAGTGGAAACAGACTGCGTGTCTAT 2367
Db 2281 GGAGGGAAGCCCTCTCTCTGGGATTTATACGATGGAGTGGAAACAGACTGCGTGTCTAT 2340
QY 2368 TTAATATGAAGCAAGTTGATCTCTCAGCCACAGCAGCCAGCGTGGTGGAGCT 2427
Db 2341 TTAATATGAAGCAAGTTGATCTCTCAGCCACAGCAGCCAGCGTGGTGGAGCT 2400
QY 2428 GATATGCCCTGCTCTGGAGCTGTTGAAGTGAACATGCAGACATGGGCTCTGTCTGT 2487
Db 2401 GATATGCCCTGCTCTGGAGCTGTTGAAGTGAACATGCAGACATGGGCTCTGTCTGT 2460
QY 2488 GATCTGATTTCTCTCTCATGCTGCAATGCTGTGAGAGAAATTAATTTGGAGAT 2547
Db 2461 GATCTGATTTCTCTCTCATGCTGCAATGCTGTGAGAGAAATTAATTTGGAGAT 2520
QY 2548 GCCATATCTCTTCTGTGGGAGATCACTTTGGAAAGGGAATGCTTAACCTTGGGCCGAA 2607
Db 2521 GCCATATCTTCTGTGGGAGATCACTTTGGAAAGGGAATGCTTAACCTTGGGCCGAA 2580
QY 2608 AAGTTCAGTGTGAAGGGAGTGAACCTCACTTTCGATTTATGCCCATTTGTTCAACATCCG 2667
Db 2581 AAGTTCAGTGTGAAGGGAGTGAACCTCACTTTCGATTTATGCCCATTTGTTCAACATCCG 2640
QY 2668 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGCTGTTCCTCCGATATACAGATGTC 2727
Db 2641 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGCTGTTCCTCCGATATACAGATGTC 2700
QY 2728 CGACTTGTGAATGCAATCCAGTGTGAGGCAAGTGGAGATCAAGCTGTGGACAC 2787
Db 2701 CGACTTGTGAATGCAATCCAGTGTGAGGCAAGTGGAGATCAAGCTGTGGACAC 2760
QY 2788 TGGGGTCACTGTGTGACACCCACTGGGACCCAGAGAAGTGGCGTGTCTTATTCAGACAG 2847
Db 2761 TGGGGTCACTGTGTGACACCCACTGGGACCCAGAGAAGTGGCGTGTCTTATTCAGACAG 2820
QY 2848 CTCAGCTGTGGGACTGCTCTCTCAACACAGGAGGAAATATATTGGAGAAAGTGT 2907
Db 2821 CTCAGCTGTGGGACTGCTCTCTCAACACAGGAGGAAATATATTGGAGAAAGTGT 2880
QY 2908 CGTGTGGGGACACAGTTTCACTTCTAGGGAATGAGTCACTTCTGGATACTCTCAA 2967
|||||

Db 2881 CBTGTGTGGGGACACAGGTTTCACTTCTAGGAAATGAGTCACTTCTGGATAACTGTCAA 2940
QY 2968 ATGACAGTTCTTGGAGCACTCCCTGTATCCATGGAATACTGTCTCTGTGTGATCTGCACA 3027
Db 2941 ATGACAGTTCTTGGAGCACTCCCTGTATCCATGGAATACTGTCTCTGTGTGATCTGCACA 3000
QY 3028 GGAAGCCTGACCCAGCCACTGTTTCCATGCTCGGAAATGATATCTGACCCATATTTGTCT 3087
Db 3001 GGAAGCCTGACCCAGCCACTGTTTCCATGCTCGGAAATGATATCTGACCCATATTTGTCT 3060
QY 3088 GCAGTTCCAGAGGGCAGTCTTGTGATCTCTTAGAGCAAAACGGCTCCGCTAGTGGAT 3147
Db 3061 GCAGTTCCAGAGGGCAGTCTTGTGATCTCTTAGAGCAAAACGGCTCCGCTAGTGGAT 3120
QY 3148 GGGGACAGCGCTGTGCCGGGAGATGATCTATCAGACAGGCTTCTGGGACACCATC 3207
Db 3121 GGGGACAGCGCTGTGCCGGGAGATGATCTATCAGACAGGCTTCTGGGACACCATC 3180
QY 3208 TGTGATGACGGCTGGGACCTGAGCGATGCCACGCTGGTGTGTCAAAAGCTGGGCTGTGGA 3267
Db 3181 TGTGATGACGGCTGGGACCTGAGCGATGCCACGCTGGTGTGTCAAAAGCTGGGCTGTGGA 3240
QY 3268 GTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGAGGGGTCAGGGCCATCTGGCTG 3327
Db 3241 GTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGAGGGGTCAGGGCCATCTGGCTG 3300
QY 3328 GATGACCTGAACCTGCACAGGAACGGAGTCCCACTTTGTCAGTGCCTTTCCCGCGGTG 3387
Db 3301 GATGACCTGAACCTGCACAGGAACGGAGTCCCACTTTGTCAGTGCCTTTCCCGCGGTG 3360
QY 3388 GGGCAGCAGCTGCAGGACACAGGAGCGCAGGGGTCTATCTGTCAGAAATTCAGACC 3447
Db 3361 GGGCAGCAGCTGCAGGACACAGGAGCGCAGGGGTCTATCTGTCAGAAATTCAGACC 3420
QY 3448 TTGAGCTCTACAGTCAAACTGAAACAGAGAGCTGTCTGGGAGATTGGAAGTCTTCTAT 3507
Db 3421 TTGAGCTCTACAGTCAAACTGAAACAGAGAGCTGTCTGGGAGATTGGAAGTCTTCTAT 3480
QY 3508 AACGGACCTTGGGCGAGCTGCGCAGGAGAAACATCACCACAGCCATAGCAGCATTTGT 3567
Db 3481 AACGGACCTTGGGCGAGCTGCGCAGGAGAAACATCACCACAGCCATAGCAGCATTTGT 3540
QY 3568 TCCAGCAGCTGGGCTGTGGGAGAAATGAGTGTCTAGCTCCGCTTATCTAGACA 3627
Db 3541 TGACGAGCTGGGCTGTGGGAGAAATGAGTGTCTAGCTCCGCTTATCTAGACA 3600
QY 3628 GGCTCTGGTTTCACTGTGGTGGATGACATTCAGTCTCTTAAACGCAATATCTCCATATGG 3687
Db 3601 GGCTCTGGTTTCACTGTGGTGGATGACATTCAGTCTCTTAAACGCAATATCTCCATATGG 3660
QY 3688 CAGTGCCTGTCTGCCCATGGGAGCGAAGAAATCTCCAGCCCGCAGAGAGACTGTGATC 3747
Db 3661 CAGTGCCTGTCTGCCCATGGGAGCGAAGAAATCTCCAGCCCGCAGAGAGACTGTGATC 3720
QY 3748 ACATGTGAGATAGATAGAGTGGTGGAGAGACACCCAGTGTCTGGAGAGTGGAG 3807
Db 3721 ACATGTGAGATAGATAGAGTGGTGGAGAGACACCCAGTGTCTGGAGAGTGGAG 3780
QY 3808 ATCTGACAGCAGGCTCTGGGACAGTGTGTGATGATCTCTGGGACCTGGCGGAGCG 3867
Db 3781 ATCTGACAGCAGGCTCTGGGACAGTGTGTGATGATCTCTGGGACCTGGCGGAGCG 3840
QY 3868 GAAAGTGTGTGACAGCTGGGCTGTGGCTGTGCTCTGCTGCTGAGGGAGCTGTG 3927
Db 3841 GAAAGTGTGTGACAGCTGGGCTGTGGCTGTGCTCTGCTGCTGAGGAGCTGTG 3900
QY 3928 TTTGGCCAGGAACTGGAACCATCTGGTTGGATGACATCGGGTGCAAGGAAATGAGTCA 3987
Db 3901 TTTGGCCAGGAACTGGAACCATCTGGTTGGATGACATCGGGTGCAAGGAAATGAGTCA 3960
QY 3988 TTTCTATGGACATGTCTCAGCCAAACCTTGGGAGCAGAGTGTGTGACACAAAGGAGAT 4047
Db 3961 TTTCTATGGACATGTCTCAGCCAAACCTTGGGAGCAGAGTGTGTGACACAAAGGAGAT 4020

QY 4048 GCTGGCGTAGGTCCTCGGACAGTCGCTGAATCACTGAATGCTCCTCAGGTCATTTA 4107
|||||
Db 4021 GCTGGCGTAGGTCCTCGGACAGTCGCTGAATCACTGAATGCTCCTCAGGTCATTTA 4080
|||||
QY 4108 GCACCTATTATTCAGCATATCTTTGGGCTCCTCTCTCGGTCCTGTTTATTTCTC 4167
|||||
Db 4081 GCACCTATTATTCAGCATATCTTTGGGCTCCTCTCTCGGTCCTGTTTATTTCTC 4140
|||||
QY 4168 ACGTGGTGGCAGGTCAGAAACAAACATCTGCCCTCAGAGTTTCAACCAGAGGAGG 4227
|||||
Db 4141 ACGTGGTGGCAGGTCAGAAACAAACATCTGCCCTCAGAGTTTCAACCAGAGGAGG 4200
|||||
QY 4228 GGTTCCTCGAGGAGAAATTTATTCATGAGATGGAGACCTGCCCTCAAGAGAGAGACCCA 4287
|||||
Db 4201 GGTTCCTCGAGGAGAAATTTATTCATGAGATGGAGACCTGCCCTCAAGAGAGAGACCCA 4260
|||||
QY 4288 CATGGGACAAGAACTCAGATGACACCCCAACCACATGGTGTGAAGATGCTAGCGACACA 4347
|||||
Db 4261 CATGGGACAAGAACTCAGATGACACCCCAACCACATGGTGTGAAGATGCTAGCGACACA 4320
|||||
QY 4348 TCGCTGTGGGAGTCTCTCTGCTCTGAAGCCACAAA 4386
|||||
Db 4321 TCGCTGTGGGAGTCTCTCTGCTCTGAAGCCACAAA 4359
|||||
RESULT 5
US-10-042-431-79
; Sequence 79, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; FILE REFERENCE: 10147-602
; CURRENT APPLICATION NUMBER: US/10/042,431
; PRIOR FILING DATE: 2001-10-25
; PRIOR FILING DATE: 1999-06-14
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 79
; LENGTH: 4308
; TYPE: DNA
; ORGANISM: Bos sp.
US-10-042-431-79

Query Match 19.28; Score 886.8; DB 9; Length 4308;
Best Local Similarity 54.98; Pred. No. 5.5e-233;
Matches 2061; Conservative 0; Mismatches 1527; Indels 168; Gaps 9;

QY 483 TCTGGGTTTGGAGTGGTGAACAACTCCTCTGTCAGGAGAGTGGAGTGAATTT 542
|||||
Db 75 TCTGGAGCTGAGTTGAAGATGAGTCCATCGCTGTGAGGGAGAGTGGAGTGAACCA 134
|||||
QY 543 CCAAGAAAGTGGGGACTATATGTGATGATGGGTGGAACATTGCTGCTGCGTGGT 602
|||||
Db 135 CCAAGGAGAAATGGGACAGTGGATGGTTACAGGTGGACATGAAGGATGCTGTAGT 194
|||||
QY 603 GTGAGGCAACTAGATGCTCATCTTTTATTTCTTCTGGAGTTGTTATAGCCCTGC 662
|||||
Db 195 GTGAGACAGCTGGGTGTGAGGTGCGCATTTGTTTCTGGAGGCTTATTTGGGCC 254
|||||
QY 663 TGTATGGCCCTTGGCTGGATGACATTTTATGCGAGGGAATGAGTTGGCACTCTG 722
|||||
Db 255 AGGACTGGCCCTTGGCTTTGTATGATGATGATGATGATGATGATGATGATGATGAT 314
|||||
QY 723 GAATTGACAGATCGTGGATGGGAAATCATGACTGCAGTGCACAAATGAGGATGTCACATT 782
|||||

Db 315 TGACTGTGAGCATTTCTAATATTAAGAGACTATCGTAATGATGCCTATATCATGTGCGGGA 374
|||||
QY 783 AACTTTGTTATGATAGTAGTGTATCTTGAACTAAGGCTTCTAGGTGGAACATAACCGCTGAT 842
|||||
Db 375 TGTCTGGAGTAGTCTGCTCAGGATTTG---TCCGTCTGGCTGGAGGAGTGAACCTGTCTC 431
|||||
QY 843 GGGGAGAGTAGAGCTGAAAATCCAAAGAGGTGGGGACCCGTATGCCACCAATAAGTGAA 902
|||||
Db 432 AGGCGAGTAGAAGTGCATTCTGGAGAGCTTGGATCCCACTGCTGTGATGGAACTTCAC 491
|||||
QY 903 CAATGCTGACGCTGATGCTGATGCAAGCAGCTTGGGATGGAACCCGACCTTCACTTCGC 962
|||||
Db 492 ACTTTGCCACTGCCAGATCATCTGTGCAGAGTGGGTGTGGCAAGGCTGTGTCTGCTCT 551
|||||
QY 963 TGGCTTGGCTCATTTGCACTCAGGCTGTGATGTTGATGGCTGTGATGGCTGTCTCTGCTC 1022
|||||
Db 552 GGGACATGAGGCTCTTCCAGAGTCCAGTCCCAAGCTTGGGCTGAAGAGTTCAGGTGGA 611
|||||
QY 1023 CGGTAATGAATCTTTCTTTGGGACTGCAGACATTTCCGGAACCCGTCAATTTTGACTGTCT 1082
|||||
Db 612 GGGGAGAGGCTGAGCTCTGGGTCTGCCCCAGAGTGCCTGTCCAGGGGACACGTGTCA 671
|||||
QY 1083 TCATCAAAACGATGTGTCTGTGATCTGCTCAGATGGAGCAGATTTGGAACCTGCCACTAGC 1142
|||||
Db 672 CCACAGTGGATCTGCTCAGGTTGTTTCTTACATCTACTCAGA---AGTCCGGCTCATGAC 728
|||||
QY 1143 AGATGGAAGTAACAATTTCTTCCAGGAGAGTAGAGTGAGATTCATGAACAGTGTGGAC 1202
|||||
Db 729 AACGGCTCTCTCAGTGTGAAGGGCAGGTGGAGATGAACATTTCTTGGACAAATGAGAGC 788
|||||
QY 1203 AATATGTACCAAGCAACTGGAAGAATGAACAGCCCTTGTGCTTTGTAAAGCAGCTAGGATG 1262
|||||
Db 789 GCTCTGTGCTCTCCACTGGAGTCTGGCCAATGCCAATGTTATCTGCTCAGCTCGGCTG 848
|||||
QY 1263 TCCGTTGACGCTCTTTGGCAGTCTGTGCTGTGCTGTAAACCTAGTAAAGCTAGAGACATTTG 1322
|||||
Db 849 TGGAGTTGCCATCTCCACCCCGGAGGACACACTTGGTGAAGAAAGGTGATCAGATCCT 908
|||||
QY 1323 GATAAAGCAGCATATCTTGCACCTGGGAATGAGTCAAGCTCTCTGGGACTGCACATATGATGG 1382
|||||
Db 909 AACAGCCGATTTCACTGCTCTGGGCTGAGTCTTCTGTTGGAGTTCCTCTGTGACTGTC 968
|||||
QY 1383 AAAAGCAAAAGCAAGATGCTTCCGAAGATCAGATGCTGGAGTAAATTTGCTGTGATAAGC 1442
|||||
Db 969 CCTGGTGGTCTCTGACTGTTCCCATGGCAACACAGCCCTCTGTGATCTCTCAGGAACCA 1028
|||||
QY 1443 AGATCTGG----- 1450
|||||
Db 1029 GATCCAGGTGCTTCCCAAGTCAACGACTCCGCTGTCTCAACCTACAGGCTCTGCGGCTC 1088
|||||
QY 1451 -----ACCTAAGGCTTGTGCGGCTCATAG 1475
|||||
Db 1089 AGAGACAGCCCTCTACTCTCAGACAGCAGGAGCTCCGCTGTGTGAGAGGGGGCGG 1148
|||||
QY 1476 CCCCTGTTATGGAGATTTGGAGTGAAATACCAAGAGAGTGGGGAGTGTGTCTGATCA 1535
|||||
Db 1149 TCCCTGGCCGGAGAGTGGAGATCTTTGACAGGCTCTCTGGGACCACTCTGTGATGA 1208
|||||
QY 1536 CAGATGGAGCACAGGAATGAGCTGTTGTGTAAACAATTTGGATGTTGGAACCCCTAT 1595
|||||
Db 1209 GGGTGGGACCTTGGACGATGCCCGCTGTGTGAGGAGCTGGGCTGTGAGAGAGCCCT 1268
|||||
QY 1596 GCATGTGTTGCTGATGACCTATTATAAGAAAGCATCAGAGCTATTTCGGCTGGATGAGCT 1655
|||||
Db 1269 CAATGCCACGGGCTCTGCTCACTTCGGGGCAGGATTCAGGGGCCCATCTGTTGGACA 1328
|||||
QY 1656 TTCTTGCAATTTGGAATGAGTCAAAATATCTGGGACTGTGAACACAGTGTGATGGGAAAGCA 1715
|||||
Db 1329 GAACGTGCACAGAAAGGAGTCCACGCTGTGAGGTGCCCTTCCCGGGGCTGGGGACCA 1388
|||||
QY 1716 TAATTGTGTACACAGAGAGGATGTGATTTAACTGTCTCAGGTGATGCAACATGGGGCT 1775
|||||

Db 1389 CAACATGCAGACAAAGCAGGAGCGGGGGTCACTCTGCTAGAGT-----TCCTGGCCCT 1442
Qy 1776 GAGGCTGTGGCGGCAGCAACCGCTGCTCGGGAAGACTGTGAGGTGTACTTTCAAGAGCG 1835
Db 1443 CAGGATGGTGAAGTGAAGGACCAAGTGTGCTGGTGGTGAAGTTTCTACAAATGGGAC 1502
Qy 1836 GTGGGCACAGTGTGTGATGACGGCTGGAAACAGTAAAGCTGCAGCTGTGGTGTGAGCCA 1895
Db 1503 CTGGGCAGTGTCTGCGGTAAACCCCAAGTGAAGACATCACTGTGTCCAGATCTGCAGACA 1562
Qy 1896 GCTGACATGCCATCTTCTATCATTTGGCATGGGTCTGGAAACGCTTCTACAGGATATGG 1955
Db 1563 GCTTGGCTGTGGGACAGTGAACCTCAACTCTTCTGTCTGTAGAGAAGTTTATG 1622
Qy 1956 AAAAATTTGGCTGCATGATTTCTGTGATGAGATGAGTCAAGTCTCTGTGATGAG 2015
Db 1623 GCCAGTGGGTAGTAGAATCCAGTGTGCGAAACAGTACACTCTCTGCGAGTGTCC 1682
Qy 2016 GAACAGTGGGTGGGAAATTAATGACTGCAGTCAAGTCAAGATGTGGAGTGTATGTTTC 2075
Db 1683 TTCTGACCCCTTGAATTAACAATCACTATGCTCTCCAAAGGAGGAGCCTATATCTGGTG--- 1739
Qy 2076 TGATGATCGGATATGGAGTGAAGTGTGGTGGTGAACACAGAGTGTGCTGGAAGT 2135
Db 1740 ---TGACAGACAGACAGATCCGCTGTGGTGGTGAAGTGTGCTGTGGAGAGT 1796
Qy 2136 TGAGGTGAATGCCAGGTGCGGTGGAAATCTGTGTCTAATGTGGTGGGAAATGAACAT 2195
Db 1797 GGAGATCTTGACAGGGCTCTGGGACCACTCTGTGATGACCTGGGACCTGACGA 1856
Qy 2196 TGGTGAAGTGTGAGCAACTTGAATGTGGGTCTGCAATCAAGGTCTCAGAGAGCC 2255
Db 1857 TGCCCGTGTGGTGTGAAGCAGTGGGCTGTGGAGAGCCCTGGAGCCACTGTCTTTC 1916
Qy 2256 TCATTTACAGAAAGAACTACACATCTTAATGTGCAATTTCTGGCTGCAGTGGAGGGA 2315
Db 1917 CTCTTTCGGGACGGATAGGCCCCACTGTGGTGGATGAAGTGAAGTCAAGAGGAGAGA 1976
Qy 2316 AGCCTCTCTGCGATGTATACGATGGAGTGGAAACAGACTGCGTGTCTATTAATAT 2375
Db 1977 GTCCCAAGTATGAGGTGCCCTTCTGGGGATGGGCGCAACACAACATGCAATCATCAAGA 2036
Qy 2376 GGAAGCAAGTTGTATGCTCAGCCACAGGACGCCAGCTGGTGGAGCTGATATGCC 2435
Db 2037 AGATGAGGAGTCACTGTCTAGGATTTGTC-----GTCTGGCTGGAGGATGGACC 2090
Qy 2436 CTGCTCTGAGCTGTGTAAGTGAACATCAGACATCGCGCTCTGTCTGTGATTCTCA 2495
Db 2091 CTGCTCAGGGCAGTGAAGTGAATCTTGAGAAGCTTGACCCCAAGTGTCTGTGGAAGA 2150
Qy 2496 TTTCTCTCTTCACTGCCAATGTGCTGCAGAGAAATTAATTTGTGGAGATGCCATATC 2555
Db 2151 CTTCACACTCCCACTGCCAGGTCTATCTGCAGAGCTGGGATGGCAAGGCTGTGTC 2210
Qy 2556 TCTTTCTGTGGAGATCACTTTGGAAGGAAGTGTCTAACTTGGGCGCAAAAGTTCCA 2615
Db 2211 TGTCTTGGGACACATGCCATTCAGAGAGTCCGATGGCCAGGCTGTGGCTCAAGAGTTCA 2270
Qy 2616 GTCTGAAGGAGTGAACATCACTTCATATGCCCATTTGTTCAACATCCGGAAGACAC 2675
Db 2271 GTGTGATGGGGGAGCCCTGAGTCTGTGGCTGCCCAAGTGCCTGTCCAGAGGACAC 2330
Qy 2676 TTGTATCCACAGCAGAGAAGTTGGATTTGTCTGTTTCCCGATATACAGATGTCCGACTGT 2735
Db 2331 ATCTCTCCACAGTGGAGCTCTCAGGTGTCTGTTTCAAGTGTACACAGAAGTCCAGCTTAT 2390
Qy 2736 G---AATGGCAATCCAGTGTGACGGCAAGTGGAGATCAAGCTGTGGACACTGGGG 2792
Db 2391 GAAAAACGGCACTCTCAATGTGAGGGCAGGTGGAGATGAAGATCTCTGGACGATGGAG 2450
Qy 2793 CTCACTGTGTGACACCACTGGGACCCAGAGATGCCGTGTCTATGACAGAGCTCAG 2852
Db 2451 AGCGCTCTGTGGCTCCCACTGGAGTGTGGCAATGCCAATGTGTCTGTCTGTGCTCGG 2510

Qy 2853 CTGTGGGACTCTCTCTCAACACAGAGAGAAATATATTGGAGAAAGAGTGTTCGTGT 2912
Db 2511 CTGTGGAGTGCCTACTCCACCCACAGAGACCACACTTGTGTGGAGAGGTGATCAGAT 2570
Qy 2913 GTGGGCACACAGGTTCATTGCTTAGGAATGAGTCACTTCTGGGATAACTGTCAATGAC 2972
Db 2571 CTCAACACCCCAATTTCACTGCTCAGGGCTGAGTCTTCTCTGGAGTGTTCCTGTGAC 2630
Qy 2973 AGTTCTTGAGACACCTCCCTGTATCCATGGAAATACTGTCTGTGATCTGCACAGGAAG 3032
Db 2631 TGCTTGGTGGGCTGACTGTTCCTATGGCAACAGCCTCTGTGATCTGCTCAGGAAA 2690
Qy 3033 COTGACCCAGCACCATGTTTCCATGCTCGCAATATATCTGACCCATATTTGTGTGCAAGT 3092
Db 2691 CCACACCCAGGTGCTGCCCACTGCAACGACTTCTGTCTCAACCTGACGGCTCTCGCGC 2750
Qy 3093 TCCAGAGGCGAGTGTCTGATCTGTAGAGGACAAACGGTCCCGCTAGTGGATGGGA 3152
Db 2751 CTGAGAGGAGTTCCTCCTACTGCTCAGACAGCAGGAGCTCCCGCTGGTGGAGCGGG 2810
Qy 3153 CAGCGCTGTGCGGAGAGTAGAGTCTATCACCAGCGCTTCTGGGCAACCATCTGTGA 3212
Db 2811 CGTCTCTCGCGGAGAGTGGAGTCTTGACCAAGGCTCTCTGGGCAACCATCTGTGA 2870
Qy 3213 TGACGGCTGGGACCTGAGCGATGCCAGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGC 3272
Db 2871 TGATGACTGGGACCTGGAGTGCCTGTGTGTGTGTCAGAGCAGCTGGGCTGTGGAGAGC 2930
Qy 3273 CTTCAATGCCAGGCTCTGCTCACTTTGGGAGGGGTGAGGCCCATCTGGCTGGATCA 3332
Db 2931 CTTCAATGCCAGGCTCTGCTCACTTTGGGAGGGGTGAGGCCCATCTGGCTGGATCA 2990
Qy 3333 CTTGAACCTGCACAGGAAGGAGTCCACTTGTGGCAGTGCCTTCCCGGGCTGGGGCA 3392
Db 2991 CTTGAACCTGCACAGGAAGGAGTCCACTTGTGGAGTGCCTTCCCGGGCTGGGGCG 3050
Qy 3393 GCACACTGCAGGACACAGGAGGAGCGAGGTGTCTGCTCAGAAATTCACAGCCTTGAG 3452
Db 3051 GCACACTGCAGACACAGGAGGAGCGGGGTCTCTGCTCAGAGTTCTCTGGCCCTCAG 3110
Qy 3453 GCTCTACTAGTGAACATGAAACAGAGAGTGTCTGCGGAGATTGGAAGTCTCTATAACGG 3512
Db 3111 G-----ATGTTGAGGAGGAGGACCACTGCTGGTGGCTGGAGGTTTCTACACGG 3164
Qy 3513 GACTGGGCGAGCTGCGGAGGAGAAACATCACACAGCCATAGCAGGCAATGTGTGACAG 3572
Db 3165 GACTGGGCGAGTCTGCGGAGCCCATGGAAGATATCACTGTCTCGTGTATCTGCAG 3224
Qy 3573 GCAGCTGGGCTGTGGGAGAAATGGAGTTTGTAGCCTTCCCGCTTTATCTAAGACAGGCTC 3632
Db 3225 ACAGCTTGGATGTGGGACAGTGAAGTCTCAACACTCTGTTGGTCTCAGGGAAGGTTTC 3284
Qy 3633 TGGTTTCATGTGGTGCATGACATTCAGTGTCTTAAACGCATATCTCCATATGGCAGTG 3692
Db 3285 TAGACCCCGTGGGTAGATTTAATTCAGTGTGCGAAATATGGATACCTCTCTCTGGCAGTG 3344
Qy 3693 COTGCTGCCCCATGGGAGGAGAAATCTCCAGCCCGCAGAGAGACCTGGATCACATG 3752
Db 3345 TCTTCTGCCCCATGAAATACAGTTCATGCTCTCCAAAGGAGGAGCCTACATCTCATG 3404
Qy 3753 TGAAGATAGATAA----- 3766
Db 3405 TGAAGAGAGAGAACCAAGAGCTGTCCAACTGCTCGCGCTGCACAGACAGAGAGAGCT 3464
Qy 3767 ---GAGTGTCTGGAGGACACCGAGTGTCTGGAGGTGGAGTCTGCGACGAGGCTC 3824
Db 3465 CCGCTCTCAGGGGAGAGACAGCGAGTGTCTCAGGGGTGGAGGTGTGCACAAACGGCTC 3524
Qy 3825 CTGGGCGACAGTGTGTGATGACTCTCTGGACCTTGGCCGAGGAGGAGTGTGTGTGACGA 3884
Db 3525 CTGGGCGACCGTGTGGATGACTCTCTGGAGCCTGGCAGAGGCTGAGGTGTGTGTGACGA 3584

Db 1149 TCCCTGCGCCGGAGTGAGATCCTTGACCAGGGCTCTTGGGCAACATCTGTGATGA 1208
Qy 1536 CAGATGGAGCACAAAGAAATGACAGCTGTTGTGTAAACAATTTGGATGTGGAAAGCCAT 1595
Db 1209 CGCTGGGACCTTGGAGATGCCCGGTGTGTGACAGGAGCTGGGCTGTGGAGAAGCCCT 1268
Qy 1596 GCATGTGTTTGGTATGACCTATTTAAAGAAGCATCAGGACCTATTTGGCTGTGATGAGCT 1655
Db 1269 CAATGCCAGGGGTGCTGCTCACTTGGGGCAGGATCAGGGCCCATCTGGTTGGACAACCT 1328
Qy 1656 TTCTGCAATGGAATGAGTCAAAATCTGGGACGTGGAACAGTGGATGGGAAGCA 1715
Db 1329 GAATGTGACAGAGAGATGTGATTTAACTCTGCTCAGGTGATGAACATGGGGCT 1775
Qy 1716 TAAATGTGTACAGAGAGATGTGATTTAACTCTGCTCAGGTGATGAACATGGGGCT 1775
Db 1389 CAACTGCAGACACAGCAGGACGCGGGGTCTCTCTCAGAGT-----TCTGSGCCCT 1442
Qy 1776 GAGGCTGTGGGCGGAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGAGC 1835
Db 1443 CAGGATGGTGTGAGTGAGGACAGCAGTGTGCTGGGTGGCTGGAAGTTTCTACAATGGGAC 1502
Qy 1836 GTGGGSCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCA 1895
Db 1503 CTGGGCGAGTGTCTGCCGTAAACCCATGGAACACATCACTGTGCCAGATCTGCAGACA 1562
Qy 1896 GCTGGAATCCCATCTCTTATCAATTTGGCATGGGTCTGGGAAACGGCTTCTACAGGATATGG 1955
Db 1563 GCTTGCTGTGGGACAGTGAACCTCAACTCTTCTGCTTTAGAGAAGTTTATAG 1622
Qy 1956 AAAAATTTGGCTGATGATGTTTCTGTGTATGGAGATGAGTCAGATCTCTGTGTCATGCA 2015
Db 1623 GCCACAGTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1682
Qy 2016 GAACAGTGGGTGGGAAATATCACTGCAGTCACAGTGAAGATGTTGGAGTATCTGTTTC 2075
Db 1683 TTCTGACCTTGGAAATCAACTATGCTCTCCAAAGGAGAGCCTATATCTGGTG---- 1739
Qy 2076 TGATGCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCTGGAAGGT 2135
Db 1740 ---TGCAGACAGACAGATCCGCTGTGGTGTGGATGAGGTGTGCTGTCTGGGAGCT 1796
Qy 2136 TGAGTGAATGCCAGGTGCGTGGGAATTTCTGTGTATGTTGCTGATGTTGGGGAATGAACAT 2195
Db 1797 GGAGATTCCTTGACAGGGCTCTGGGGCCACCATCTGTGATGACCGCTGGGACCTGGACGA 1856
Qy 2196 TGTGAAGTTGTTTGAGGCAACTTCAATGTGGTCTGCAATCAGGCTTCCAGAGAGCC 2255
Db 1857 TGCCGTGTGGTGTGCAACAGCTGGGCTGTGGAGAGCCCTGGACGCCACTGTCTCTTC 1916
Qy 2256 TCATTTACAGAAAGAACATTAACATCTTAATGTGCAATTTGGCTGCACTGGAGGGGA 2315
Db 1917 CTCTTCGGGACGGGATCAGGGCCCATCGGCTGGATGAAGTGAATGCAGAGGAGAGA 1976
Qy 2316 AGCCTCTCTCTGGGATGTATAGGATGGAGTGGAGTGGAAACAGACTCGGTGTCAATTAATAT 2375
Db 1977 GTCCCAAGTATGGAGTGGCTTCTTGGGGATGGGCGCAACAACTGCAATCATCAAGA 2036
Qy 2376 GGAAGCAAGTTTGTCTGCTCAGCCACAGGAGCCAGGCTGGTTGGAGCTGATATGCC 2435
Db 2037 AGATGAGGAGTCACTGTCTCAGGATTTGTG-----GTCTGGCTGGAGAGATGGACC 2090
Qy 2436 CTGCTCTGAGCTGTTGAAGTGAACATCAGACATAGCAGCTGCTGTCTGTGATTTCTGA 2495
Db 2091 CTGCTCAGGCGAGTAGAAGTGCATTTCTGGAGAAGCTTGACCCCAAGTGTCTGATGAAA 2150
Qy 2496 TTCTCTCTTCATCTGCGCAATGTGCTGTGCAGAGAAATTAATTTGGAGATGCCATATC 2555
Db 2151 CTTCACTACCTCCCACTGCCAGGTCTATCTGTGCAGAGCTGGGATGGCAAGGCTGTGTC 2210
Qy 2556 TCTTTCTGTGGGAGATCACTTTTGGAAAGGAATGGTCTAACTTGGGCCGCAAAAGTTCCA 2615
Db 2211 TGTCTGGGACACATGCCAATTCAGAGAGTCCGATGCCAGGCTGGGCTGAAGATTCTAG 2270

Qy 2616 GTGTGAAGGAGTGAACCTCACTTGCATATATGCCCATTTGTTCAACATCCGGAAGACAC 2675
Db 2271 GTGTGATGGGGGAGCCTGAGCTCTGTGCTGCCCCAGAGTCCCTGTCCAGGAGGCAC 2330
Qy 2676 TTGTATCCACAGCAGAGAAAGTTGGAGTTGTCTGTGTTCCCGATATACAGATGTCGACTTGT 2735
Db 2331 ATGTCTCCACAGTGGAGCTGCTCAGTTGTCTGTCTAGTGTACACAGAAATCCAGCTTAT 2390
Qy 2736 G---AATGSCAAATCCAGTGTGACGGCAAGTGGAGATCAACGTCTTGGACACTGGGG 2792
Db 2391 GAAAAACGGCAGCTCTCAATGTGAGGGGAGGTGGAGATGAAGATCTCTGGACATGGAG 2450
Qy 2793 CTCACTGTGTACACCACTGGGACCCAGAGATGCCGTGTCTATGACAGACAGCTCAG 2852
Db 2451 AGCGCTGTGTGCTCCACTGCCAGTCTGGCAATGCAATGTTGTCTGTCTGCTGAGCTCG 2510
Qy 2853 CTGTGGGACTGCTCTCAACACAGGAGGAAATATATTTGGAGAAAGTGTTCCTGTCT 2912
Db 2511 CTGTGAGTGGCCATCTCCACCCAGAGGACCACTTGGTGGAGGAGGTGATCAGAT 2570
Qy 2913 GTGGGACACAGGTTTCAATTTAGGGAATGAGTCACCTTCTGGATAAATGTCAAAATGAC 2972
Db 2571 CTCACAGCCCAATTTCACTGCTCAGGGCTGAGTCTTCTCTGGAGTTGTCTCTGAC 2630
Qy 2973 AGTTCTTGGAGCACCTTCCCTGTATCCATGGAATACTGTCTGTGATCTGCACAGGAAG 3032
Db 2631 TGCTTGGGTGGGCTGACTTCCCATGGCAACACAGGCTCTGTGATCTGTCTCAGGAAA 2690
Qy 3033 CCTGACCCAGCACGTTTCCATGCTCCCAATGTATCTGACCCATATTTGTCTCTCAGT 3092
Db 2691 CCACACCCAGGTGCTGCCCCAGTGCAACGACTTCTGTCTCAACCTGACAGGCTCTCGGG 2750
Qy 3093 TCCAGAGGACGTGCTTTGATCTGTAGAGCAAAACGGCTCCGCTAGTGGATGGGA 3152
Db 2751 CTCAGAGAGATTTCTCCCTACTGCTCAGACAGCAGGAGCTCCGCTGGTGGACGGGG 2810
Qy 3153 CAGCCGCTGTGCGGGAGAGTAGAGATCTATCAGCAGCGCTTCTGGGGACCACTCTGTGA 3212
Db 2811 CGTCTCCCTGCGGGGAGAGTGGAGATCTTGAACAGGCTCTCTGGGGACCACTCTGTGA 2870
Qy 3213 TGACGCTGGGACCTCAGCGATGCCACCTGTGTCTCAAAAGCTGGGCTGTGGAGTGGC 3272
Db 2871 TGATGACTGGGACCTGGAGATGCCGCTGTGTGTGTCAGGACAGTGGGCTGTGGAGAGC 2930
Qy 3273 CTTCAATGCAAGGCTCTCTGCTCACTTTTGGGAGGGGTGAGGGCCCATCTGGCTGTGATGA 3332
Db 2931 CCTCAATGCCAGGGTCTGCTCACTTCGSGGAGGATCAGGSCCATCTGGCTGGACGA 2990
Qy 3333 CCTGAATGCACAGGAACGGAGTCCCACTTTGGGAGTGGCTTCCCGGGCTGGGGGCA 3392
Db 2991 CCTGAATGCACAGGAAGAGTCCCACTGTGGAGTGGCTTCCCGGGCTGGGGGGG 3050
Qy 3393 GCAGACTGCAGGCAACAGGAGGACGAGGCTCTCTCAGAAATTCACAGCTTGTAG 3452
Db 3051 GCAGACTGCAGACACAGAGAGGACCGGGGTCTCTCTCAGAGTTCTGTGGCCCTCAG 3110
Qy 3453 GCTCTACAGTGAACATGAACAGAGAGCTGTGTGGAGATTGGAAAGTCTTCTATAACGG 3512
Db 3111 G-----ATGGTGAGGAGGACCAAGCAGTGTGTGGTGGCTGGAGGTTTCTTACAACGG 3164
Qy 3513 GACCTGGGACGCTGGGAGAGGAGAAATCACCAGCCATAGAGGATGTGTGTCAG 3572
Db 3165 GACCTGGGACGTTGTGCCGAGGCCCATGGAAGATATCACTGTCTGTCTGATCTGTCAG 3224
Qy 3573 GCAGCTGGGCTGTGGGAGAAATGGAGTTGTGACGCTCGGCCCTTTATCTTAAGACAGGCTC 3632
Db 3225 ACAGCTTGTGATGGGAGACAGTGAAGTCTCAACACTCTGTGTGCTCAGGGAAGGTTTC 3284
Qy 3633 TGTCTTCTGCTGGGTGGATGACATTCAGTGTCTTAAACGCAATATCCATATGCGAGTG 3692
Db 3285 TAGACCCCGGTGGTAGATTTAATTCAGTGTGCGGAAATGGATACCTCTCTCTGTGCGAGTG 3344

Db 391 TGGATGACATCGCGTCAAGAAATGAGTCAATTTCTATGGGACTGTACGCCAACCCCT 450

QY 4016 GGGGACAGTACTGTGGACA 4037

Db 451 GGGGACAGTACTGTGGACA 472

RESULT 9

US-10-040-739-138

; Sequence 138, Application US/10040739

; Patent No. US20020173635A1

; GENERAL INFORMATION:

; APPLICANT: Jacobs, Kenneth

; McCoy, John

; LaValle, Edward

; Racie, Lisa

; Merberg, David

; Treacy, Maurice

; Spaulding, Vikki

; TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS

; NUMBER OF SEQUENCES: 1519

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: Genetics Institute, Inc.

; STREET: 87 CambridgePark Drive

; CITY: Cambridge

; STATE: Massachusetts

; COUNTRY: U.S.A.

; ZIP: 02140

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk

; COMPUTER: IBM PC Compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/040,739

; FILING DATE: 07-Jan-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/036,520

; FILING DATE: 03-JUN-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Brown, Scott A.

; REGISTRATION NUMBER: 32,724

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 498-8224

; TELEFAX: (617) 876-5851

; INFORMATION FOR SEQ ID NO: 138:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 599 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; SEQUENCE DESCRIPTION: SEQ ID NO: 138:

US-10-040-739-138

Query Match 4.6%; Score 213; DB 9; Length 599;

Best Local Similarity 68.0%; Pred. No. 4.6e-48;

Matches 297; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 1770 GGGCGTGGCTGGGGGGGCGAGCAACCGCTGCTCGGGAAGCTGGAGTCTACTTTCA 1829

Db 161 GGAGCTGAGCGTAGTGTTGTAACCAAGTGTACCGGGAGAGTGGAAAGTGAAGTCCA 220

QY 1830 AGGACGCTGGGCGACAGTGTGTATGACGCGCTGGAACAGTAAGCTGCAGCTGTGTGTG 1889

Db 221 GGAGGAGTGGGCAACGCTGTGTAATATGCTGGAGCATGGAAGCGTCTCTGTGATTG 280

QY 1890 TAGCCAGCTGGAGTCCCACTCTCTATCATTTGTCATGGTCTGGGAAAGCTTCTACAG 1949

Db 281 TAACCAAGCTGGGATGTCCCAACTGCTATCAAGCCCTGGATGGGCTTAATTCAGTGCA 340

QY 1950 ATATGAAAAAATTTGGCTCGATGATGTTTCTGTGTATGGAGATGAGTCAGATCTCTG 2009

Db 341 TTCTGGACGCAATTGGATGGATCATGTTTCTGTGGGAATGATCAGCTCTTTGGGA 400

QY 2010 ATGCAAGAACAGTGGGTGGGAAATATGACTGCAGTGCAGTCAAGTGAAGATGTTGGAGT 2069

Db 401 TTCCAAACATGATGGATGGGGAAGCATTTATGCAATCAACAATGAAGATCTGGCGTGAC 460

QY 2070 CTGTTCTGATCGCATCGGATATGAGCTGAGGCTGTGGGTGGGAAGCAGCAGGTGTGCTGG 2129

Db 461 ATGTTCTGATGGATCATGATCTGGAGCTAAGACTTAGAGGTGGAGGCGCGCTGTGCTGG 520

QY 2130 ARAAGTTGAGTGAATGTCAGGGTCCCGTGGGAATTTCTGTGCTAATGCTGGGGAAT 2189

Db 521 GACAGTTGAGTGGAGATTACAGAGACTGTTAGGAAGAGTGTGTGCACAGAGCTGGGACT 580

QY 2190 GAACATTGCTGAAGTTG 2206

Db 581 GAAAGAAGCTGAAGCTG 597

RESULT 10

US-09-796-692-269

; Sequence 269, Application US/09796692

; Publication No. US20020198362A1

; GENERAL INFORMATION:

; APPLICANT: Gaiger, Alexander

; Mannion, Jane

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER

; FILE REFERENCE: 2077,001200

; CURRENT APPLICATION NUMBER: US/09/796,692

; PRIOR FILING DATE: 2001-03-01

; PRIOR APPLICATION NUMBER: 60/186,126

; PRIOR FILING DATE: 2000-03-01

; PRIOR APPLICATION NUMBER: 60/190,479

; PRIOR FILING DATE: 2000-03-17

; PRIOR APPLICATION NUMBER: 60/200,545

; PRIOR FILING DATE: 2000-04-27

; PRIOR APPLICATION NUMBER: 60/200,303

; PRIOR FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: 60/200,779

; PRIOR FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: 60/200,999

; PRIOR FILING DATE: 2000-05-01

; PRIOR APPLICATION NUMBER: 60/202,084

; PRIOR FILING DATE: 2000-05-04

; PRIOR APPLICATION NUMBER: 60/206,201

; PRIOR FILING DATE: 2000-05-22

; PRIOR APPLICATION NUMBER: 60/218,950

; PRIOR FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: 60/222,903

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: 60/223,416

; PRIOR FILING DATE: 2000-08-04

; PRIOR APPLICATION NUMBER: 60/223,378

; PRIOR FILING DATE: 2000-08-07

; NUMBER OF SEQ ID NOS: 9597

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 269

; LENGTH: 461

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-796-692-269

Query Match 4.3%; Score 199.2; DB 9; Length 461;

Best Local Similarity 66.7%; Pred. No. 2.4e-44;

Matches 299; Conservative 0; Mismatches 148; Indels 1; Gaps 1;

QY 183 TGGAGACGGTCCCTGCTCTGGGACAGTGGAGGTGAAATCCAGGACAGTGGGGGACTCT 242

Db 5 TGGAGGGAATATGTGTTCTGGAAGATAGAGATCAATTCAGAGCGGTGGGACACT 64


```

; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 269
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-269

```

Query Match 4.3%; Score 199.2; DB 9; Length 461;

183	QY	TGAGACGGTCCCTGCTCTGGGACACTGAGGTGAAATTCACAGGACACAGTGGGGACTGT	242
5	Db	TGAGGGAATATCTGTCTTGGAGAATAGAGATCAAATTCGAAGGACGGTGGGAACAGT	64
243	QY	GTGTGATGATGGGTGGAACACTACTGCCCTCAACTGTCGTGTGCAACACACTCTGGATGTCC	302
65	Db	GTGTGATGATAACTTCACACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGG	124
303	QY	ATTTCTCTTGGCCATGTTTGGTCTTGGACAAGCCGTGACTAGACATGGAATAATTTGGCT	362
125	Db	AAGTGTGTCAAGTTCTCTGGTTCATCTAATTTGGAGAAGGCTCTGGACCAATCTGGTT	184
363	QY	TGATGATGTTTCCCTGTTA-TGGAAATGAGTCAGCTCTCTGGGNAATGTCACACCCGGGAA	421
185	Db	TGATGATCTTATATGCAACGGAAATGAGTCAGCTCTTGGAACTGCAACACTCAAGGAT	244
422	QY	GGGAAAGCCATACTGTTATCATCGAGAGAAGATGTTGGTGTGAACCTGTTATGGTGAAGCCA	481
245	Db	GGGAAAGCATAACTGTGATCATGCTGAGGATGCTGAGTGATTTGCTCAAAGGGAGCAG	304
482	QY	ATCTGGGTTTGAAGCTAGTGGATGGAAACAACCTCTCTTTCAGGAGAGTGGAGGTGAAT	541
305	Db	ATCTGAGCCTTGAGACTGGTAGATGGAGTCACTGAATGTTTCAGAGAATTAGAAGTGAGAT	364
542	QY	TCCAAGAAAGGTGGGACACTATATGTGATGTGGGTGGAACTTGAATACCTGTCGCGTGG	601
365	Db	TCCAAGGAAATGGGGACACATATGTATGATCGGCTGGACAGTTACGATGCTGCTGG	424
602	QY	TGTGACGGCAACTAGGATGTCCATCTTC	629
425	Db	CATGCAAGCAACTGGGATGTCCAACATGC	452

RESULT 13
US-10-040-862-7354
Sequence 7354, Application US/10040862
Publication No. US20030078396A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy of Hematological Malignancies
FILE REFERENCE: 014058-01352005
CURRENT APPLICATION NUMBER: US/10/040,862
CURRENT FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303

```

; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 7354
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-7354

```

Query Match
4.38; Score 199.2; DB 9; Length 461;

Qy	183	TTGGAGACGGTCCCTGCTCTGGGACAGTGGAGGTGAAATTCACGGGACAGTGGGGAGACTG	242
Db	5	TTGGAGGAATATGTGTCTTGGAAATAGAGATCAAAATTCGAAGGACGGTGGGAACAGT	64
Qy	243	GTGTGATGATGGGTGGAACTACTGCCCTCAACTGCTGTCGACAAACGCTGGATGTCC	302
Db	65	GTGTGATGATAACTTCACATAGATCATGCTCTGTCATTTGTAGACAACCTTGAATGTGG	124
Qy	303	ATTTCTTTCGCCATGTTTCGTTTTTGGACAAGCCGTGACTAGACATGGAATAATTTGGCT	362
Db	125	AAGTGTCTGACGTTCTCTGGTTCATTAATTTGGAGAAGGCTCTGGACCAATCTGGTT	184
Qy	363	TGATGATGTTTCTCTGTTA-TGGAAATGAGTCAGCTCTCTGGGAATCTCAACACCGGGAAT	421
Db	185	TGATGATCTTATATGCAACCGGAATAGATGACGCTCTTGGAACTGCAACACATCAAGAT	244
Qy	422	GGGGAACCAATACTGTTATCATCGGAGAAGATGTTGGTGTGAACCTGTTATGGTGAAGCCA	481
Db	245	GGGGAAGCATAACTGTGATCATGCTGAGATGCTGAGTGATTTGCTCAAAGGGAGCAG	304
Qy	482	ATCTGGGTTTGAGGCTAGTGGATGGAAACAACTCCTGTTCAAGGAGAGTGGAGGTGAAT	541
Db	305	ATCTGAGCCTTGAGACTGGTAGATGGAGTCACCTGAATGTTTCAGGAAGATTAGAAGTGAGAT	364
Qy	542	TCCAAGAAAGCTGGGGACTATATGTGATGATGGGTGGAACCTTGAATACTGTCGCCGTGG	601
Db	365	TCCAGGAGAAATGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGG	424
Qy	602	TGTGCAGGCAACTAGGATGTCCATCTTC	629
Db	425	CATGCAAGCAACTGGGATGTCCAACTGC	452

RESULT 14
US-09-147-947-5
; Sequence 5, Application US/09147947A
; Patent No. US20020160490A1
; GENERAL INFORMATION:
; APPLICANT: TSURUOKA, NO. US20020160490A1uo
; APPLICANT: YAMASHIRO, Kyoto
; APPLICANT: YAMAGUCHI, NO. US20020160490A1omi
; TITLE OF INVENTION: NO. US20020160490A1el
; protease

FILE REFERENCE: 001560-349
CURRENT APPLICATION NUMBER: US/09/147,947A
CURRENT FILING DATE: 1997-03-24
EARLIER APPLICATION NUMBER: PCT/JP98/03324
EARLIER FILING DATE: 1998-07-24
EARLIER APPLICATION NUMBER: JP 9/213969
EARLIER FILING DATE: 1997-07-24
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 2562
TYPE: DNA
ORGANISM: Human
FEATURE:
OTHER INFORMATION:
US-09-147-947-5

Query Match 4.3%; Score 197.6; DB 9; Length 2562;
Best Local Similarity 52.2%; Pred. No. 2.1e-43;
Matches 505; Conservative 0; Mismatches 439; Indels 24; Gaps 2;

Qy 1133 TGCAGCTAGCAGATGGAAGTAACAATTTTCAGGAGAGTAGAGGTGAGAAATTCATGAAC 1192
Db 580 TTCGCCCTTGTGGAGGAGCAGCTGTGCATGAGCCGGTGGAGCTCTACCATGCTGCC 739
Qy 1193 AGTGGTGGACAATATGTGACAGAACTTGAAGAAATGAACAGCCCTTGTGTTTGAAGC 1252
Db 740 AGTGGGAACCGTTTGTGTGACCAATGGATGATGCCGATGAGAGTGTCTGCGAGC 799
Qy 1253 AGCTAGATGTCCTTTCAGGCTCTTTGGCAGTCTGTGCTGTAAACCTAGTAAGTGAAGCTA 1312
Db 800 AGCTGGGCTCAGTGGCATTTGCCAAGCATGGCATAGGCATATATTTGGGAAGGCTG 859
Qy 1313 GAGACATTTGGATAACAGCATATCTTGCAGCTGGGAATGAGTCACTCTCTGGGACTGCA 1372
Db 860 GCCCAGTTATTTGGATGAGTACGCTGCACTGGGAATGAGCTTTCAATTGAGCAGTGT 919
Qy 1373 CATATGATGAAAAAGCAAGCAATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTT 1432
Db 920 CAAGAGCTCTCGGAGAGCATAACTGTGCCATAAAGAGATGCTGGAGTCTCTGTA 979
Qy 1433 CTGATAGGCGAGATCTGGACCTAAGGCTTGTGGGGCTCATAGCCCTGTTATGGGAGAT 1492
Db 980 CCCCTCTAAACAGATGGGTCTATCAGACTTGCAGGTGGGAAGGAGCAGCATGAGGCTGCT 1039
Qy 1493 TGGAGGTGAATACCAAGGAGTGGGAGTGTCTCATGACAGATGGAGCACAAGGA 1552
Db 1040 TGGAGGTATATTACAGAGCCAGTGGGAAGTGTCTGTGATGATGGCTGGACTGAGCTGA 1099
Qy 1553 ATGCAGCTGTTGTGTAAACAATTTGGGATGTGGAAGCCCTATGCATGTGTTTGGTATGA 1612
Db 1100 ATACATAGCTGTTGTGCGAGTGGGATTTAAATATGTTAAACAAGCATCTG--CCA 1156
Qy 1613 CCTATTTTAAAGACATCAGGACCTATTTGGCTGGATGAGCTTCTTTCATGGAATG 1672
Db 1157 ACCATTTTGAAGAAGCAGAGGCGCCATATGTTGGATGAGCTGAGCTGCTCAGGAAAG 1216
Qy 1673 AGTCAATATCTGGAGCTGTGAACACAGTGGATGGGAAGCAATATGTTGTACACAGAG 1732
Db 1217 AAACAGATTTCTCAGTGTGTTCCAGGCGACAGTGGGAAGGATGACTGCGACCCAGCG 1276
Qy 1733 AGGATGTGATTGTAACTCTCAGGTGATGCAACATGGGGCTG----- 1776
Db 1277 AAGATGTTAGCATTTGCCTGTCTACCTCGCGCGAGGAGACAGGCTCTCTCTGGGTTTC 1336
Qy 1777 -----AGCTGGTGGCGGAGCAACCCCTGCTCGGGAAGACTGGAGGTGTACTTCAAG 1831
Db 1337 CTGTCAGCATGATGGATGGAGAAAATAAGAAAGAAGCAGAGTGGAGTGTATTAATCAATG 1396
Qy 1832 GACGGTGGGCGACAGTGTGTGATGACGGCTGGGAACAGTAAAGCTGAGCTGTGTGTGTA 1891
Db 1397 GCCAGTGGGAACAATCTGTGATGATGGATGATGAAGGATGCAGCTGTGTATCTGTC 1456

Qy 1892 GCCAGCTGGAGTGCCTCATCTTCTATCATTTGGCATGGGTCTGGGAAACGCTTCTACAGAT 1951
Db 1457 GTACAGTTGGCTACAAGGGTCTCCAGAGCAACAACATGGCTTACTTTGGAGAAGAA 1516
Qy 1952 ATGGAAAAATTTGGCTCGATGATGTTTCTCTGTGATGAGAGTGTCTGAGATCTCTGGTCA 2011
Db 1517 AAGGACCCATCCATGTGGAATAATGTGAAGTGCACAGGAATGAGAGGTCCTTGGCTGACT 1576
Qy 2012 GCAGGAACAGTGGTGGGGAATAATGACTGCAGTGCAGTGAAGATGTTTGGAGTGTAT 2071
Db 1577 GTATCAAGCAGATATTGGGAAGACACAACTGCCCCACAGATGAAGATGCAGAGTTATTT 1636
Qy 2072 GTTCTGAT 2079
Db 1637 GTGATTAT 1644

RESULT 15
US-09-147-947-3
Sequence 3, Application US/09147947A
Patent No. US20020160490A1
GENERAL INFORMATION:
APPLICANT: TSURUOKA, No. US20020160490A1uo
APPLICANT: YAMASHIRO, Kyoto
APPLICANT: YAMAGUCHI, No. US20020160490A1omi
TITLE OF INVENTION: No. US20020160490A1el Serine Protease
FILE REFERENCE: 001560-349
CURRENT APPLICATION NUMBER: US/09/147,947A
CURRENT FILING DATE: 1997-03-24
EARLIER APPLICATION NUMBER: PCT/JP98/03324
EARLIER FILING DATE: 1998-07-24
EARLIER APPLICATION NUMBER: JP 9/213969
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 2614
TYPE: DNA
ORGANISM: Mouse
FEATURE:
OTHER INFORMATION:
US-09-147-947-3

Query Match 4.2%; Score 193.2; DB 9; Length 2614;
Best Local Similarity 52.0%; Pred. No. 3.5e-42;
Matches 500; Conservative 0; Mismatches 438; Indels 24; Gaps 2;

Qy 1133 TGCAGCTAGCAGATGGAAGTAACAATTTTTCAGGAGAGTAGAGGTGAGAAATTCATGAAC 1192
Db 741 TTCGCCCTTGTGGTGGGAACAGTGGGCATGAAGGTGCGAGTGGAGCTGTACCACGCTGCC 800
Qy 1193 AGTGGTGGACAATATGTCACAGAACTGGAAGATGAACAGCCCTTGTGTTTGTAGC 1252
Db 801 AGTGGGAGACCATCTGTGACGACCAATGGGACATGAGAGCAGACAGCTCATCTGTAGC 860
Qy 1253 AGCTAGGATGTCCTTTCAGGCTCTTTGGCAGTCTGCTGCTGCTTAAACCTAGTAAGAGCTA 1312
Db 861 AGCTGGGCTCAGTGGCATTTGCCAAGCATGGCATGCGACATTTTGGGAGAGTCTG 920
Qy 1313 GAGACATTTGGATAAAGCAGCATATCTTGCATGGAATGAGTCACTGCTCTCTGGAGTGA 1372
Db 921 GCCCAATATTGTTGATGAAGTACGCTGCAACGGAACAGCTGTCAATTGAGCAATGTC 980
Qy 1373 CATATGATGGAAGCAAGCAACATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTT 1432
Db 981 CAAGAGTCTCTGGGGCGCAACATAAATCTGTGCCATAAAGAAAGATGCTGGAGTGTCTTGTG 1040
Qy 1433 CTGATAAGGAGATGCTGACCTTAAGGCTTGTGCGGGCTCATAGCCCTGTTTATGGGAGAT 1492
Db 1041 TTCCTCTAAACAGATGGTGTCTATCAGACTGGCAGGAGGAAAAGTACCATGAGGTCGCC 1100
Qy 1493 TGGAGGTGAATACCAAGGAGAGTGGGGAGTGTGTGTCTATGACAGATGGAGCACAAGGA 1552
Db 1552 TGGAGGTGAATACCAAGGAGAGTGGGGAGTGTGTGTCTATGACAGATGGAGCACAAGGA 1552

Db 1101 TGGAGGCTCTACTACAAAGGGGAGCTGGGGGACAGCTGTGTGATGATGGCTGGAGTGAGATGA 1160
QY 1553 ATGCAGCTGTTGTGTAAACAATTGGGATGTGGAAGCCTATGCATGTGTTTGGTATGA 1612
Db 1161 ACACATACGTGCTTGTGACTGCTGGGATTTAAATACGGCAACA---GTCCTCTGTGA 1217
QY 1613 CTTATTTTAAAGAAGCATCAGACCTATTGGCTGGATGACGTTTCTTGCATTGGAATG 1672
Db 1218 ACCATTTTGTATGGCAGCAACAGGCCCATATGGCTGGATGACCTGCTCAGGAAAAG 1277
QY 1673 AGTCAATATCTGGGACTGTGAACACAGTGGATGGGAAAGCATAATTGTGTACACAGAG 1732
Db 1278 AGTCAGCTTTCATTGAGTGTCCAGGAGACAGTGGGGAAGGCATGACTGCACCCATAGAG 1337
QY 1733 AGGATGTGATTGTAACTGCTCAGGTGA-----TGCAACATGGG 1771
Db 1338 AAGATGTGGGCTCACCTGCTATCCTGACAGCGATGGACATAGGCTTTCTCCAGGTTTC 1397
QY 1772 GCCTGAGGCTGTGGCGGCAGCAACCCGCTGCTCGGGAAGACTGGAGGTGTACTTCAAG 1831
Db 1398 CCATCAGACTAGTGGATGGAGAGAATAAGAAGGAAGGAGGAGTGGAGGTTTTTGTCAATG 1457
QY 1832 GACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGA 1891
Db 1458 GCCATGGGGACAACTCTCGGATGACGGATGGACCGATGAAGCATGCAGCTGTGATCGCC 1517
QY 1892 GCCAGCTGGACTGCCCATCTTCTATCATTTGGCATGGGCTGGGAAACGCTTCTACAGGAT 1951
Db 1518 GGCAGCTTGGCTATAAGGCTCTGCCAGAGCAAGGACTATGGCTTATTTTGGGGAAGGAA 1577
QY 1952 ATGGAAAAATTTGGCTCGATGATGTTCTCTGTGTGGAGATGAGTCAGATCTCTGGTCAT 2011
Db 1578 AAGGCCCATCCACATGGGATAATGTGAAGTGCACAGGAAATGAGAAGGCCCTGGCTGACT 1637
QY 2012 GCAGGAACAGTGGTGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTCATCT 2071
Db 1638 GTGTCAACAAGACATTGGAAGGCACAACTGCCGCCACAGTGGAGTGCAGGATCATCT 1697
QY 2072 GT 2073
Db 1698 GT 1699

Search completed: May 12, 2003, 06:36:37
Job time : 658 secs

Db 1546 CTTGCTGTCATGCACAGTAGTAGGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGAGG 1605
Qy 1794 CAACCGCTCTCGGAGAGACTGGAGGTGTACTTTCAAGGACCGTGGGCGACAGTGTGA 1853
Db 1606 TGACAGGTGTGAGGCGGAGTGGAGGTCTATACCAAGGCTCTCTGGGCGACCGTGTGGGA 1665
Qy 1854 TGACGGCTGGAACAGTAAGCTGACAGCTGTGGTGTAGCCAGCTGGACTGGCCATCTTC 1913
Db 1666 TGACAGCTGGGACACCAATGATGCCATGTCTCTGTCAGGCAACCGGCTGTGGCTGGGC 1725
Qy 1914 TATCATTTGATGGTCTGGGAAACGCTTCTACAGGATATGAAATAATTTGGCTCGATGA 1973
Db 1726 CATGTGAGCCAGGAAATGCCCGGTTTGGTCAGGCTCAGGACCAATGTCTTGATGA 1785
Qy 1974 TGTTCCTGTGATGGAGATGATGATCTCTGGTCATGCAGGACAGTGGGTGGGAAA 2033
Db 1786 TGTGGCTGCTCTCAGGACACAGGTCTTACCCGTGGAGCTGCCCCACAAATGGCTGGCTC 1845
Qy 2034 TAATGACTGAGTCACAGTGAAGATTTGGAGTGTCTCTTCTGATGC----- 2081
Db 1846 CCACACTGTGGCCATGTAAGACGCTGGTGTCTCTCAGCTTCCAGTCCCGGCC 1905
Qy 2082 -----ATCGGA 2087
Db 1906 AACACCTAGTCCAGACACTTGGCCAACTTCATGATCATCAACGACGAGTCTGAATCCAG 1965
Qy 2088 TATGGAGCTGAGGCTTGTGGGTGGAAGCAGAGGTGTGCTGGAATAATTTGAGTGAATCT 2147
Db 1966 TTTGGCCCTGAGGCTGTGTAATGAGAGTGCACAGGTGTGAGGCCGAGTGGAGTCTCTATA 2025
Qy 2148 CCAGGCTGGCTGGGAATTTCTGTGCTAATGGCTGGGAAATGAACATTTGCTGAATTT 2207
Db 2026 CCAGGCTCTCTGGGACCGCTGTGATGACTACTTGGGACACCAATGATGCCAATGTGT 2085
Qy 2208 TTCAGGCAACTGATGTGGTCTGCATCAGGCTCTCCAGAGCCTCATTTACACA 2267
Db 2086 TTCAGGAGCTGGGCTGTGGCTGGGCCATGTGACGCCAGGAAATGCCCCGTTTGGCCA 2145
Qy 2268 AAGAACATTACATCTTAATGTCGAATTTGGCTGCACTGGAGGGGAAAGCTCTCTCTG 2327
Db 2146 GGTTCAGGACCAATTTGCTGTGATGATGTGCTCTCAGGACATGAGTCTATCTGTG 2205
Qy 2328 GGAATGTATACGATGGGAGTGGAAACAGAGTGGCTGTCTATTAATATGAGCAAGTTT 2387
Db 2206 GAGCTGCCCCACAATGGCTGGCTCTCCACAACTGTGSCATCATGAAGACGCTGTGT 2265
Qy 2388 GATCTGCT-----CAGCCACAGGACGCC----- 2412
Db 2266 CATCTGCTAGCTTCCAGTCCAGCCGACACCCAGCCAGACACTTGGCCAACTCACA 2325
Qy 2413 -----AGGTGGTTGGAGCTGATATGCC 2435
Db 2326 TGCATCAACAGCAGGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGAGTGACAG 2385
Qy 2436 CTGCTCTGGAGCTGTGAAGTGAACATGCAGACATGGCGCTCTGTCTGATTTCTGA 2495
Db 2386 GTGTGAGGCGGAGTGAGGTCTATACCGAGGCTCTGGGCGACCGTGTGTGATGACATA 2445
Qy 2496 TTTCTCTCTTCTGCTGCCAATGTGCTGCAGAGAAATTAATTTGGAGATGCCATATC 2555
Db 2446 CTGGACACCAATGATGCCAATGTGTTGACAGCAGCTGGCTGTGGTGGGCCAGCTC 2505
Qy 2556 TCTTTCTGTGGGAGATCACTTTGAAAGGAAATGTTTAACTTGGGCGGAAAGTTCCA 2615
Db 2506 AGCCCCAGGAAATGCCCGGTTTGGCCAGGTTTTCAGGACCAATTTGCTGATGATGTGG 2565
Qy 2616 GTGTGAGGAGTCAAACTCACTTGCATTTATGCCCATTTGTTCAACATCCGGAAGACAC 2675
Db 2566 CTGCTCAGGACATGAGTCTCTATCTGTGGAGCTGCCGCCCAATGGCTGGCTCTCCACAA 2625
Qy 2676 TTGTATCCACAGCAGCAAGTGTGAGTGTCTGTCTTCCGAT----- 2716
Db 2626 CTGTGGCCATCATGAAGACGCTGGTGTCTCTGCTCAGCTTCCAGTCCAGCGGACAC 2685

Qy 2717 -----ATACAGA 2723
Db 2686 CAGCCGACACTTGGCCAACTCTCGTGCATCAACAGCAGGATCTGAATCCACTTTGGC 2745
Qy 2724 TGTCCGACTGTGAATGCAAAATCCCACTGTGACGGGCAAGTGGAGATCAACGCTTTGG 2783
Db 2746 CTTGAGACTGTGAATGGAGGTGACAGGTGCGAGGCCGAGTGGAGGTTCTTATACAAAG 2805
Qy 2784 ACATCTGGGCTCAGTGTGTGACACCACTGGGACCCACAAAGATGCCCTGTCTTCTATGAG 2843
Db 2806 CTCTGGGGCACCGTGTGTGATGACTTGGGACCAATGATGCCAAGCTGTGCTGAG 2865
Qy 2844 ACAGCTCAGCTGTGGGACTGTCTCTCAACACAGGAGGAAATATATTGGAGAAAGAG 2903
Db 2866 CGAGCTGGGCTGTGGCTGGCCATGTGAGCCCAAGTATGCCAGTTTGGCCAGGCTC 2925
Qy 2904 TGTTCGTGTGGGACACAGGTTTCAATTTGCTTAGGAAATGAGTCTCTGGAATCTG 2963
Db 2926 AGGACCCATTTCTCTGGATGTGCTGCTCAGGACACGAGTCTTACCTGTGGAGCTG 2985
Qy 2964 TCAATGACAGTTCTTGGAGCACCTCCCTGTATCATGGAATACTGTCTGTGATCTG 3023
Db 2986 CCCCACAAATGGCTGGCTCTCCACAACTGTGGCCATCATGAGATGCTGTGCTATCTG 3045
Qy 3024 CACAGGAAGCTTGACCCAGCCACTGTTTCCATGCTCGCAATGTATCTGACCCATATT 3083
Db 3046 CTCAGCTGTCTCAGTCCCAAGTCAAG-----CCAGGCCAGATACTTGGCTGACCACTT 3102
Qy 3084 GTCTCAGTTCCAGAGGCGAGTGTGATCTGTCTAGGACAAACGGCTCCGCTAGT 3143
Db 3103 ACCGCTGATGACATGAGTATGATCTGAATCCAGTT-----TGGCTCTGAGGCTGT 3150
Qy 3144 GGATGGGACAGCCGCTGTGCGGAGAGTAGAGATCTATCAGCAGCGCTTCTGGGGCAC 3203
Db 3151 GAATGGAGTGCAGGTGTGAGGCCGAGTGCAGGCTCTGTATCAGGCTCTCTGGGAAAC 3210
Qy 3204 CATCTGTGATGACGGCTGGACCTGAGCGATGCCACGTTGGTGTGTCAAAAGCTGGGCTG 3263
Db 3211 CGTGTGTGATGACAGTGGGACACCAATGATGCCAATGTGCTGTCAGGACGCTGGGCTG 3270
Qy 3264 TGGAGTGGCTTCAATGCCACGCTCTCTGCTCAGTTTGGGAGGGTCAAGGCCCATCTG 3323
Db 3271 TGGCTGGGCGATGTGCGCCAGGAATGCCGGTTTGGCCAGGCTCAGGACCACTTGT 3330
Qy 3324 GCTGATGACCTGAACTGCACAGGACGAGTCCCACTTGTGCGAGTGCCTTCCGCGG 3383
Db 3331 CTTGATGATGTGCGCTGCTCAGGGAATGAGTCTACCTGTGGAGTGTGCTGCTGCTG 3390
Qy 3384 CTGGGGGACGAGCTGAGGACACAGGACGAGGAGGCTGAGGCTGCTGCTGCTGAG 3436
Db 3391 CTGGCTCAGCCCAACTGTGGCCATCATCAAGAGCGCTGGTGTCTGCTGCTGAG 3443

RESULT 2

US-08-470-350B-1

; Sequence 1, Application US/08470350B

; Patent No. 5684126

; GENERAL INFORMATION:

; APPLICANT: Li, Xiao

; APPLICANT: Snyder, Solomon H

; TITLE OF INVENTION: Ebnerin: A Secreted von Ebner's Gland

; TITLE OF INVENTION: Protein Associated with Taste Buds

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Banner & Witcoff, Ltd.

; STREET: 1001 G Street, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20001

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

```

QY 3289 TCTGCTCACTTTTGGGAGGGGTCAAGGCCCCATCTTGCTGGATGACCTGAACCTGCACAGGA 3348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1123 AGTGGCCACTTTGGACAAGGCTCTGGATCCATTGTTCTGGATGACGTGGCCTGTACAGGA 1182
QY 3349 ACGGAGTGCCACTTTGGCAGTGGCCCTTCCCGCGGCTGGGGGACACGACACTGCAGGCAC 3408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1183 CATGAGGCCATCTGTGGAGCTGCTCCCAAGGCTGGCTCTCTCATAACTGTGGCCAC 1242
QY 3409 AAGGAGGACGACGGGTCTATCTGCTCAGAAATTCACAGCCTTTCAG 3452
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1243 CATGAGGATGCTGGATCATCTGTTTCAGATGCCCAACCCAGAG 1286

RESULT 3
US-09-341-587-6
; Sequence 6, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341,587
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: PCT/DE98/00096
; EARLIER FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 29598
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-341-587-6

Query Match 3.0%; Score 140.8; DB 4; Length 29598;
Best Local Similarity 66.4%; Pred. No. 1.2e-28;
Matches 202; Conservative 0; Mismatches 102; Indels 0; Gaps

QY 3133 CTCGCGCTAGTGATGGGACAGCCGCTGTGCCGGGAGAGTAGAGATCTATCAGCAGCGC 3192
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7049 CTGAGGCTGTGTAATGGAGCGACAGGTGTGAGGCCGAGTGGAGTCCCTATACCGAGGC 7108
QY 3193 TTCTGGGGACACATCTGTGATGACGCTGGGACCTGAGGGATGCCACAGCTGGTGTGTCAA 3252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7109 TCCTGGGGCACCCGTGTGTGATGACTACTGGGACACCAATGATGCCAATGTGCTGTGAGG 7168
QY 3253 AAGCTGGGCTGTGAGTGGCCCTTCAATGCCACGGTCTGTGCTCACTTTGGGAGGGGTCA 3312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7169 CAGCTGGGCTGTGGCTGGGCCATGTCAGGCCCAAGAAATGCCAGTTTGGCCAGGGCTCA 7288
QY 3313 CGGCCCATCTGGCTGGATGACCTGAACTGCACAGAAACGGAGTCCCACTTGTGGCAGTGC 3372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7229 GGACCCATTTCTCTGGATGATGCGCTCTCTCAGGACACGAGTCTCTACTCTGTGGAGTGC 7288
QY 3373 CTTTCCCGGGCTGGGGCAGCAGCATGCAGGCACAAAGGAGGACGCGGGGTCACTTCG 3432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7289 CCCACAAATGGCTGGCTCAACCCACAACTGTGCCCATAGTGAAGAGCGCTGGTGTCTATCGC 7348
QY 3433 TCAG 3436
    ||| |||
Db 7349 TCAG 7352

RESULT 4
US-09-341-587-7
; Sequence 7, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341,587
; CURRENT FILING DATE: 1999-08-31

```


EARLIER APPLICATION NUMBER: PCT/DE98/00096
EARLIER FILING DATE: 1998-01-09
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 28720
TYPE: DNA
ORGANISM: Homo sapiens
US-09-341-587-7

Query Match 3.0%; Score 139.2; DB 4; Length 28720;
Best Local Similarity 66.1%; Pred. No. 3.4e-28;
Matches 201; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 3133 CTCGCGCTAGTGGGACAGCCGCTGTGCGGGAGAGTAGATCTATCAGCAGCGC 3192
DB 2944 CTGAGACTGGTGAATGGAGGTGACAGGTGTGCGAGCCGAGTGGAGTCTATACCAAGC 3003
QY 3193 TTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCACGTGGTGTCTCAA 3252
DB 3004 TCCTGGGGCACCCTGTGTGATGACTACTGGGACACCAATGATGCCAACGTGCTGCAGG 3063
QY 3253 AAGCTGGGCTGGAGTGGCTTCAATGCCACGGTCTCTGCTCACTTTGGGAGGGTCA 3312
DB 3064 CAGCTGGGCTGTGGCTGGGCCATGTGACGCCAGGAAATGCCAGTTTGGCCAGGGCTCA 3123
QY 3313 GGGCCCATCTGGCTGATGACCTGAACCTGACAGCAAGGAGTCCACCTGTGGCAGTGC 3372
DB 3124 GAACCATTTGCTCTGGATGATGTGGCTGCTCAGACAGAGTCTTACCTGTGGAGCTGC 3183
QY 3373 CTTTCCCGCGGTGGGGCAGCAGCACTGACGAGCAAGGAGGAGCGAGGGTCTATCTGC 3432
DB 3184 CCCCACAATGGCTGCTCCCAACAACCTGTGCCATCATGAAGATGCTGTGTCTATCTGC 3243
QY 3433 TCAG 3436
DB 3244 TCAG 3247

RESULT 5
US-09-341-587-2
Sequence 2, Application US/09341587
Patent No. 6346606
GENERAL INFORMATION:
APPLICANT: Mollenhauer, Jan
TITLE OF INVENTION: Protein Containing an SRCR Domain
FILE REFERENCE: 4121-108
CURRENT APPLICATION NUMBER: US/09/341,587
CURRENT FILING DATE: 1999-08-31
EARLIER APPLICATION NUMBER: PCT/DE98/00096
EARLIER FILING DATE: 1998-01-09
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2001
TYPE: DNA
ORGANISM: Homo sapiens
US-09-341-587-2

Query Match 2.9%; Score 136; DB 4; Length 2001;
Best Local Similarity 65.5%; Pred. No. 5.1e-28;
Matches 199; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 3133 CTCGCGCTAGTGGGACAGCCGCTGTGCGGGAGAGTAGATCTATCAGCAGCGC 3192
DB 184 CTGAGCTGGTGAATGGAGGTGACAGGTGTGCGAGCCGAGTGGAGTCTGTATCAGAGC 243
QY 3193 TTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCACGTGGTGTCTCAA 3252
DB 244 TCCTGGGAACCCCTGTGTGATGACAGCTGGGACACCAATGATGCCAATGTGCTCTGCAGG 303
QY 3253 AAGCTGGGCTGGAGTGGCTTCAATGCCACGGTCTCTGCTCACTTTTGGGAGGGTCA 3312
|||||

DB 304 CAGCTGGGCTGTGGCTGGGCCATGTGCGCCCAAGGAAATGCCGGTTTGGCCAGGCTCA 363
QY 3313 GGGCCCATCTGGCTGGATGACCTGAACCTGACAGCAAGCAGAGTCCACACTTGTGGCAGTGC 3372
DB 364 GGACCATTTGCTCTGGATGATGTGGCTGCTCAGGAATGAGTCTACTCTGTGGAGCTGC 423
QY 3373 CTTTCCCGCGGTGGGGCAGCAGCACTGACGAGCAAGGAGGAGCGAGGGTCTATCTGC 3432
DB 424 CCCCACAAGGCTGGCTCACCCACAACACTGTGCCATCATCAAGAGCAGCTGTGTCTATCTGC 483
QY 3433 TCAG 3436
DB 484 TCAG 487

RESULT 6
US-09-034-916-1
Sequence 1, Application US/09034916
Patent No. 6046314
GENERAL INFORMATION:
APPLICANT: GEBE, JOHN A.
APPLICANT: SIADAK, ANTHONY W.
APPLICANT: ARUFFO, ALEJANDRO A.
TITLE OF INVENTION: SPALPHA: A NOVEL SCAVENGER RECEPTOR
TITLE OF INVENTION: CYSTEINE-RICH DOMAIN-CONTAINING POLYPEPTIDE AND MONOCLONAL
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROBINS & ASSOCIATES
STREET: 90 MIDDLEFIELD ROAD, SUITE 200
CITY: MENLO PARK
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/034,916
FILING DATE: 04-MAR-1998
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/039,956
FILING DATE: 06-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 5998-0020
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 325-7812
TELEFAX: (650) 325-7823
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2178 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 64..1101
US-09-034-916-1

Query Match 2.8%; Score 127.6; DB 3; Length 2178;
Best Local Similarity 51.1%; Pred. No. 1.3e-25;
Matches 499; Conservative 0; Mismatches 439; Indels 38; Gaps 7;
QY 3134 TCCTGCTAGTGGGAGCAGCGCTGTGCGGGAGAGTAGAGATCTATCAGCAGCGCT 3193
DB 131 TCGCGCTGTGGGGGCGCTCCACCGCTGTGAAGGGCGGCTGGAGGTGGAACAAAGGCC 190
|||||

Qy 3194 TCTGGGCGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCACGCTGGTGTGTCAAA 3253
Dy 191 AGTGGGCGCACCATCTGTGATGACGGCTGGGACCTAAGGACGCTGGTGTGTGTGTGGCGG 250
Qy 3254 AGCTGGGCTGTGGAGTGGCCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTGAG 3313
Dy 251 AGCTGGGCTGTGGAGTGGCCCGGAGGACCCCTAGTGGTATTTGTATGAGCCACGAGAG 310
Qy 3314 GGCCCATCTGGCTGATGACCT-GAAGTGCACAGGAGGAGTCCCACTTCTGGCGAGTGC 3372
Dy 311 AAAAGAGCAAAAGGTCTCTATCAATCAGTCAGTTGCAGAGGACAGAAATATATGG 370
Qy 3373 CTTTCCCGGCTGGGGGAG--CACGACTGCAGCAAGAGGAGGAGGAGGAGGAGGAGTCACT 3430
Dy 371 CTCAGTGTGAGCAAGAAGTATGATTTTATGATTTTATGATGATGATGATGATGATGATG 429
Qy 3431 GCTCAGAAATTCACAGCTTGGAGTCTACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3472
Dy 430 TGTGAGAACCCAGAGAGCTCTTTCTCCCGAGTCCCGAGGAGTGTGAGGCTGGCTGACGGC 489
Qy 3473 -CAGAGAGCTGTGCTGGGAGATTTGAAGTCTTCTATAACGGGACCTGGGGGAGGCTGGC 3531
Dy 490 CTTGGGCAATTCAGGGGAGGCTGGAAGTGAAGCAGGAGGAGGAGGAGGAGGAGGAGGAG 549
Qy 3532 AGGAGGAATATCACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3591
Dy 550 CAGACAGGCTGGAGCTCCGGGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 609
Qy 3592 AATGAGT---TGTGAGCTGCCCCCTTATCTAAGACAGGCTGTGTTTATGATGAGGAGT 3648
Dy 610 GCTGTACTGACTCAAAACGCTGCAACAGGAGTGCCTATGGCCGAGGAGGAGGAGGAGGAG 669
Qy 3649 GATGACATTCAGTGTCTTAAAGGAGTATCTCCATGAGGAGTGTGCTGCTGCCCATAGG 3708
Dy 670 AGCCAGATGTATGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 729
Qy 3709 GAGCGAAGAAATCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3762
Dy 730 GGAAGAGACACTGCAACCATGATGAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 789
Qy 3763 ATAAGAGTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3822
Dy 790 TTGAGACTAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 849
Qy 3823 TCCTGGGCGACAGTGTGATGAGTCTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3882
Dy 850 GTATGGGCTCTGTCTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 909
Qy 3883 CAGCTGGGCTGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3936
Dy 910 CAACCTGGGCTGTGGAGGAGTCCCT 969
Qy 3937 GGAAGTGAACCAATCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3996
Dy 970 GGGGTGTGGGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1029
Qy 3997 CACTGTACGCGCAACCTCTGGGAGGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4056
Dy 1030 CAGTGCAGAGCAGATTTTGGGGTTTTCAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1089
Qy 4057 AGGTGCTCTGGAGCT 4072
Dy 1090 ATCTGCTCAGGATAGT 1105

RESULT 7

US-08-477-674-9

; Sequence 9, Application US/08477674

; Patent No. 5644035

; GENERAL INFORMATION:

; APPLICANT: Kothe, Kirston E.

; APPLICANT: Halenbeck, Robert F.

; APPLICANT: Taylor, Eric W.

APPLICANT: Wang, Alice M.
APPLICANT: Caspitt, Clayton L.
TITLE OF INVENTION: Secreted Mac-2-Binding Glycoprotein
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cetus Oncology Corporation
STREET: 1400 Fifty-Third Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/477,674
APPLICATION NUMBER: US/08/477,674
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/316,714
FILING DATE:
APPLICATION NUMBER: US/07/961,404
FILING DATE: 15-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Kenneth M.
REGISTRATION NUMBER: 34,174
REFERENCE/DOCKET NUMBER: 2595.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 420-3152
TELEFAX: (510) 658-5470
TELEX: N/A
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2285 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-477-674-9

Query Match

2.6%; Score 118.6; DB 1; Length 2285;

Best Local Similarity 62.1%; Pred. No. 4.4e-23;

Matches 187; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Qy 3134 TCCGCTAGTGGATGGGACACCGCTGTGCGGGAGAGTAGAGATCTATCAGGCGCT 3193
Dy 250 TCGGCTGGCCGATGGGGCGCCACCAACGAGGCGCGGTGGAGATCTTCTACAGAGCC 309
Qy 3194 TCTGGGCGACCATCTGTGATGACGCTGGGACCTGAGCGATCCCGACGCTGGTGTGTCAAA 3253
Dy 310 AGTGGGCGACTGTGTGACAACTGTGGGACCTGACTGATGCCAGCGCTGCTGCCGGG 369
Qy 3254 AGCTGGGCTGTGGAGTGGCCCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTGAG 3313
Dy 370 CCCTGGGCTTTCAGAAACGCCACCGAGGCTCTGGGAGAGCTGCCCTTGGGCAAGGATGAG 429
Qy 3314 GGCCCATCTGGCTGATGACCTGAAGTGCACAGGAGGAGTCCCACTTGTGGCAGTGC 3373
Dy 430 CCCCCATCATGCTGGACGAGGTCCAGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 489
Qy 3374 CTTCCCGGCGCTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3433
Dy 490 AGTCCCTGGCTGGCTGAAGACCACTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 549
Qy 3434 C 3434
Dy 550 C 550

RESULT 8

US-08-473-791-9

```
; Sequence 9, Application US/08473791
; Patent No. 5736340
; GENERAL INFORMATION:
; APPLICANT: Koths, Kirston E.
; APPLICANT: Halenbeck, Robert F.
; APPLICANT: Taylor, Eric W.
; APPLICANT: Wang, Alice M.
; APPLICANT: Casipit, Clayton L.
; TITLE OF INVENTION: Secreted Mac-2-Binding Glycoprotein
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cetus Oncology Corporation
; STREET: 1400 Fifty-Third Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,791
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/316,714
; FILING DATE:
; APPLICATION NUMBER: US/07/961,404
; FILING DATE: 15-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Kenneth M.
; REGISTRATION NUMBER: 34,174
; REFERENCE/DOCKET NUMBER: 2595.1
; TELEPHONE: (510) 420-3152
; TELEFAX: (510) 658-5470
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2285 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-473-791-9

Query Match 2.6%; Score 118.6; DB 1; Length 2285;
Best Local Similarity 62.1%; Pred. No. 4.4e-23;
Matches 187; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 3134 TCCGCTAGTGGATGGGACAGCGCTGTGCCGGGAGAGTAGAGATCTATCAGACGGCT 3193
Db 250 TGGCGCTGGCGATGGGGGCGCCACCAACAGGCGCGTGGAGATCTTCTACAGAGGCC 309
QY 3194 TCTGGGACCACTGTGTGATGACGGCTGGGACCTGAGCGATGCCACGTTGTGTCAA 3253
Db 310 AGTGGGCACTGTGTGTGACAACCTGTGGGACCTGACTGATGCCAGGCTGTCTGCCGGG 369
QY 3254 AGCTGGGCTGTGGAGTGGCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGTTCAG 3313
Db 370 CCCTGGGCTTCGAGAACGCCACCGAGCTCTGGGACAGCTGCTTCCGGAAGGATCAG 429
QY 3314 GGCCCATCTGGCTGGATGACCTGAACAGGAGTCCCACTTTGGGAGTGCC 3373
Db 430 GCCCATCATCTGGAGAGGTCAGTGCACGGGAACCGAGGCTCTACTGGCCAGTGA 489
QY 3374 CTTCCCGGGCTGGGGGACGACGACTGCAGGACACAGGAGGAGGAGGCTATCTGCT 3433
Db 490 AGTCCCTGGGCTGTGTAAGAGCAACTGCAGGACAGAGAGAGAGCGCTGGTGTGCTGCA 549
QY 3434 C 3434
```

```
; Sequence 9, Application US/08316714
; Patent No. 5965382
; GENERAL INFORMATION:
; APPLICANT: Koths, Kirston E.
; APPLICANT: Halenbeck, Robert F.
; APPLICANT: Taylor, Eric W.
; APPLICANT: Wang, Alice M.
; APPLICANT: Casipit, Clayton L.
; TITLE OF INVENTION: Secreted Mac-2-Binding Glycoprotein
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cetus Oncology Corporation
; STREET: 1400 Fifty-Third Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/316,714
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/961,404
; FILING DATE: 15-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Kenneth M.
; REGISTRATION NUMBER: 34,174
; REFERENCE/DOCKET NUMBER: 2595.1
; TELEPHONE: (510) 420-3152
; TELEFAX: (510) 658-5470
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2285 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-316-714-9

Query Match 2.6%; Score 118.6; DB 2; Length 2285;
Best Local Similarity 62.1%; Pred. No. 4.4e-23;
Matches 187; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 3134 TCCGCTAGTGGATGGGACAGCGCTGTGCCGGGAGAGTAGAGATCTATCAGACGGCT 3193
Db 250 TGGCGCTGGCGATGGGGGCGCCACCAACAGGCGCGTGGAGATCTTCTACAGAGGCC 309
QY 3194 TCTGGGACCACTGTGTGATGACGGCTGGGACCTGAGCGATGCCACGTTGTGTCAA 3253
Db 310 AGTGGGCACTGTGTGTGACAACCTGTGGGACCTGACTGATGCCAGGCTGTCTGCCGGG 369
QY 3254 AGCTGGGCTGTGGAGTGGCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGTTCAG 3313
Db 370 CCCTGGGCTTCGAGAACGCCACCGAGCTCTGGGACAGCTGCTTCCGGAAGGATCAG 429
QY 3314 GGCCCATCTGGCTGGATGACCTGAACAGGAGTCCCACTTTGGGAGTGCC 3373
Db 430 GCCCATCATCTGGAGAGGTCAGTGCACGGGAACCGAGGCTCTACTGGCCAGTGA 489
QY 3374 CTTCCCGGGCTGGGGGACGACGACTGCAGGACACAGGAGGAGGAGGCTATCTGCT 3433
Db 490 AGTCCCTGGGCTGTGTAAGAGCAACTGCAGGACAGAGAGAGAGCGCTGGTGTGCTGCA 549
QY 3434 C 3434
```

[illegible]

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2262
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-448-076-3

Query Match 2.3%; Score 108.2; DB 4; Length 2262;
Best Local Similarity 61.6%; Pred. No. 3.7e-20;
Matches 173; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 3157 CGCTGTCCGGGAGAGTAGATATACAGAGCGGCTTCTGGGGACCACTCTGTGATGAC 3216
Db 157 CCTACGAGGCGCGGTGGAGATACAGCGAGCTGTGAATGGGACCACTCTGCGATGAT 216

QY 3217 GCGTGGGACCTGAGGATGCCACAGCTGTGTCTCAAAAGCTGGGCTGTGGAGTGGCCTTC 3276
Db 217 GACTTTCACGCTCAGGCTGCCACATCTCTCCGGGAGCTGGCTTCACAGAGGCCACA 276

QY 3277 AATGCCACGGTCTCTGCTCACTTTGGGAGGGGTACAGGCCCATCTGGCTGGATGACCTG 3336
Db 277 GCGTGGACCCACAGTGCCTTGTGGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 336

QY 3337 AACTGCACAGGAGGAGTCCCACTTGTGGCAGTGCCTTCCGGGCTGGGGGAGCAGC 3396
Db 337 AGCTGCAGTGGACCGAGCAGAGTGTGACTGAATGTGCTCCCGGGCTGGGGGAGACAGT 396

QY 3397 GACTGCAGGACCAAGGAGCAGGAGGCTCATCTGTCTCAGA 3437
Db 397 GACTGTACGACGATGAGGATGCTGGGGTCACTCTGCAAGA 437

RESULT 13

US-09-702-572-3
; Sequence 3, Application US/09702572
; Patent No. 6391602

; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran

; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
; FILE OF INVENTION: USES THEREFOR

; FILE REFERENCE: MNI-073

; CURRENT APPLICATION NUMBER: US/09/702,572

; PRIOR FILING DATE: 2000-10-31

; PRIOR APPLICATION NUMBER: 09/276,400

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 2262

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-702-572-3

Query Match 2.3%; Score 108.2; DB 4; Length 2262;
Best Local Similarity 61.6%; Pred. No. 3.7e-20;
Matches 173; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 3157 CGCTGTCCGGGAGAGTAGATATACAGAGCGGCTTCTGGGGACCACTCTGTGATGAC 3216
Db 157 CCTACGAGGCGCGGTGGAGATACAGCGAGCTGTGAATGGGACCACTCTGCGATGAT 216

QY 3217 GCGTGGGACCTGAGGATGCCACAGCTGTGTCTCAAAAGCTGGGCTGTGGAGTGGCCTTC 3276
Db 217 GACTTTCACGCTCAGGCTGCCACATCTCTCCGGGAGCTGGGCTTCACAGAGGCCACA 276

QY 3277 AATGCCACGGTCTCTGCTCACTTTGGGAGGGGTACAGGCCCATCTGGCTGGATGACCTG 3336
Db 277 GCGTGGACCCACAGTGCCTTGTGGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 336

QY 3337 AACTGCACAGGAGGAGTCCCACTTGTGGCAGTGCCTTCCCGGGCTGGGGGAGCAGC 3396
Db 337 AGCTGCAGTGGACCGAGCAGAGTGTGACTGAATGTGCTCCCGGGCTGGGGGAGACAGT 396

QY 3397 GACTGCAGGACCAAGGAGCAGCGAGGGGTCACTCTGCTCAGA 3437
Db 397 GACTGTACGACGATGAGGATGCTGGGGTCACTCTGCAAGA 437

RESULT 14

US-09-276-400-1

; Sequence 1, Application US/09276400
; Patent No. 6140056

; GENERAL INFORMATION:

; APPLICANT: Khodadoust, Mehran

; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND

; FILE OF INVENTION: USES THEREFOR

; FILE REFERENCE: MNI-073

; CURRENT APPLICATION NUMBER: US/09/276,400

; CURRENT FILING DATE: 1999-03-25

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 2920

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (143)..(2401)

US-09-276-400-1

Query Match 2.3%; Score 108.2; DB 3; Length 2920;
Best Local Similarity 61.6%; Pred. No. 4.3e-20;
Matches 173; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 3157 CGCTGTCCGGGAGAGTAGATATACAGAGCGGCTTCTGGGGACCACTCTGTGATGAC 3216
Db 299 CCTACGAGGCGCGGTGGAGATACAGCGAGCTGTGAATGGGACCACTCTGCGATGAT 358

QY 3217 GCGTGGGACCTGAGGATGCCACAGCTGTGTCTCAAAAGCTGGGCTGTGGAGTGGCCTTC 3276
Db 359 GACTTTCACGCTCAGGCTGCCACATCTCTCCGGGAGCTGGGCTTCACAGAGGCCACA 418

QY 3277 AATGCCACGGTCTCTGCTCACTTTGGGAGGGGTACAGGCCCATCTGGCTGGATGACCTG 3336
Db 419 GCGTGGACCCACAGTGCCTTGTGGAGGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGG 478

QY 3337 AACTGCACAGGAGGAGTCCCACTTGTGGCAGTGCCTTCCCGGGCTGGGGGAGCAGC 3396
Db 479 AGCTGCAGTGGGAGCGGAGCAGAGTGTGACTGAATGTGCTCCCGGGCTGGGGGAGCAGT 538

QY 3397 GACTGCAGGACCAAGGAGCAGCGAGGGGTCACTCTGCTCAGA 3437
Db 539 GACTGTACGACGATGAGGATGCTGGGGTCACTCTGCAAGA 579

RESULT 15

US-09-448-076-1

; Sequence 1, Application US/09448076
; Patent No. 6300092

; GENERAL INFORMATION:

; APPLICANT: Khodadoust, Mehran et al.

; TITLE OF INVENTION: METHODS OF USE OF A NOVEL LYSYL OXIDASE-RELATED PROTEIN

; FILE OF INVENTION: MNI-073CP

; CURRENT APPLICATION NUMBER: US/09/448,076

; CURRENT FILING DATE: 1999-11-23

; EARLIER APPLICATION NUMBER: 60/117,580

; EARLIER FILING DATE: 1999-01-27

; EARLIER APPLICATION NUMBER: 09/276,400

; EARLIER FILING DATE: 1999-03-25

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 2920

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

NAME/KEY: CDS
LOCATION: (143)...(2401)
US-09-448-076-1

Query Match 2.38; Score 108.2; DB 4; Length 2920;
Best Local Similarity 61.6%; Pred. No. 4.3e-20;
Matches 173; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

Qy	3157	CGCTGTGCCGGAGAGTAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGAC	3216
Db	299	CCCTACGAGGCCGCTGGAGATACAGCGAGCTGGTGAATGGGGCACCATCTGGATGAT	358
Qy	3217	GGCTGGGACCTGAGCGATGCCCGCTGGTGTCAAAAGCTGGGCTGTGGAGTGGCCTTC	3276
Db	359	GACTTCACGCTGCAGGCTGCCACATCTCTGCCGGGAGCTGGGCTTCACAGAGGCCACA	418
Qy	3277	AATGCCACGGTCTCTGCTCACTTTGGGGAGGGTTCAGGGCCCATCTGGCTGGATGACCTG	3336
Db	419	GGCTGGACCCACAGTGCCAAATATGGCCCTGGACAGGCCCATCTGGCTGGCAACTTG	478
Qy	3337	AACTGCACAGGAACGAGTCCCACTTTGTGGCAGTGCCCTTCCCGGGCTGGGGGCAGCAC	3396
Db	479	AGCTGCAGTGGGACCGACAGAGTGTGACTGAATGTGCTCCCGGGCTGGGGGAACAGT	538
Qy	3397	GACTGCAGGCACAGGAGGACGAGGGGTATCTGCTCAGA	3437
Db	539	GACTGTACGCACGATGAGGATGCTGGGGTCACTGCAAGA	579

Search completed: May 12, 2003, 06:29:49
Job time : 374 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2003, 02:05:59 ; Search time 608 Seconds
(without alignments)
17141.856 Million cell updates/sec

Title: US-09-759-130b-379

Perfect score: 4628

Sequence: 1 gcggcgctcgcatctaga.....aaacggacgctggtggtgac 4628

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

1: /SID22/gcgdata/geneseq/geneseq-nbml/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-nbml/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-nbml/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-nbml/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-nbml/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-nbml/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-nbml/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-nbml/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-nbml/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-nbml/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-nbml/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-nbml/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-nbml/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-nbml/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-nbml/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-nbml/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-nbml/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-nbml/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-nbml/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-nbml/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-nbml/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-nbml/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-nbml/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-nbml/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4628	100.0	4628	22	AAF45123 Human TANGO 234 CD
2	4359	94.2	4359	22	AAF45124 Human TANGO 234 OR
3	3459.8	74.8	3670	22	AAS01217 DNA encoding human
4	2761.2	59.7	2800	22	AAS44609 Human full-length
5	1448.2	31.3	3716	22	AAH58649 Human polynucleoti
6	1448.2	31.3	3834	22	AAH98740 Human EST-derived
7	1448.2	31.3	3834	22	AAI60435 Human polynucleoti
8	1448.2	31.3	3834	22	AAI60436 Human polynucleoti
9	1409.4	30.5	3707	22	AAS94922 Human DNA sequence

10	1339.2	28.9	3811	22	AAI58650 Human polynucleoti
11	886.8	19.2	4308	22	AAF45141 Bovine WC1 ORF. B
12	740	16.0	820	21	AAA43017 Human secreted exp
13	642	13.9	690	22	AAH99258 Human protein enco
14	563	12.2	608	20	AAH99272 EST clone CFI18.
15	551.4	11.9	562	22	AAH99326 Human protein enco
16	447.4	9.7	449	24	ABL87362 Human ovarian canc
17	404	8.7	450	22	ABA08369 Human M160 precurs
18	379.2	8.2	527	22	AA544781 Human contig polyn
19	335.4	7.2	415	22	ABA09428 Human Cys-rich sca
20	325.8	7.0	339	24	ABN24521 Human ORFX polynuc
21	319.2	6.9	2697	24	AAS17590 DNA encoding novel
22	213.6	4.6	599	20	AAV90288 EST clone DK64. H
23	213	4.6	599	20	AAV87660 EST clone DK399.
24	207.8	4.5	702	21	AA43661 Mouse secreted exp
25	199.2	4.3	461	22	AAK54544 Human haematologic
26	197.8	4.3	5802	19	AAV49652 Human SRCR protein
27	197.6	4.3	2562	20	AAI19024 Human serine prote
28	197.6	4.3	3350	20	AAV72589 Human neurotysin
29	193.2	4.2	2614	20	AAI19027 Mouse serine prote
30	191.6	4.1	2376	20	AAV72590 Mouse neurotysin
31	173.4	3.7	444	21	AACT8302 Human cancer assoc
32	167	3.6	474	22	AAH32999 Human colon cancer
33	162	3.5	1269	22	AAH32999 Human colon cancer
34	149.6	3.2	4344	24	ABK63805 Rat sequence diffe
35	149.6	3.2	4360	18	AAI44068 Rat von Ebner's gl
36	143.8	3.1	1165	22	ABA08386 Human M130 antigen
37	143.6	3.1	1589	22	AAH33692 Human colon cancer
38	143.2	3.1	697	21	AACT99009 Human pancreatic c
39	140.8	3.0	29598	19	AAV49654 Human SC2 DNA. Ho
40	140.2	3.0	1659	22	AAI16354 Human sbg14862SPER
41	139.2	3.0	28720	19	AAV49655 Human SC3 DNA. Ho
42	136	2.9	2000	19	AAV49651 Human SRCR protein
43	134.4	2.9	713	20	AAV88878 EST clone HV435.
44	127.8	2.8	1591	19	AAV49556 Human liver cell c
45	127.2	2.7	584	22	AAH98116 Murine 7-transmemb

ALIGNMENTS

RESULT 1

AAF45123

ID AAF45123 standard; cDNA: 4628 BP.

XX AAF45123;

AC AAF45123;

XX 30-MAR-2001 (first entry)

DT Human TANGO 234 cDNA.

DE Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;

DE global nervous system; focal brain disorder; bipolar affective disorder;

DE global diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;

DE senile dementia; Huntington's disease; amyotrophic lateral sclerosis;

DE Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;

DE neuropsychiatric; psychoactive substance use; anxiety; ss.

DE Homo sapiens.

OS WO200077239-A2.

XX 21-DEC-2000.

XX 24-MAY-2000; 2000WO-US14858.

XX 14-JUN-1999; 99US-0333159.

XX (MILL-) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX WPI; 2001-032313/04.

Db 1741 ATGTAACTCTCAGGTGATGCAACATGGGGCCCTGAGGCTGGTGGCGGCGAGCAACGCC 1800
Qy 1801 TGCTCGGAAGACCTGGAGGTGCTACTTTCAAGGACGGTGGGCGACAGTGTGTGATGACGGC 1860
Db 1801 TGCTCGGAAGACCTGGAGGTGCTACTTTCAAGGACGGTGGGCGACAGTGTGTGATGACGGC 1860
Qy 1861 TGGAAACAGTAAAGTGCAGCTGTGGTGTGTAGCCAGCTGGGACTGCCCATCTTCTATCAT 1920
Db 1861 TGGAAACAGTAAAGTGCAGCTGTGGTGTGTAGCCAGCTGGGACTGCCCATCTTCTATCAT 1920
Qy 1921 GGCATGGTCTGGGAACGCTTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCC 1980
Db 1921 GGCATGGTCTGGGAACGCTTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCC 1980
Qy 1981 TGTGATGAGATGAGTCAAGTCTCTGGTCAATGAGCAACAGTGGTGGGGAATAATGAC 2040
Db 1981 TGTGATGAGATGAGTCAAGTCTCTGGTCAATGAGCAACAGTGGTGGGGAATAATGAC 2040
Qy 2041 TGCAGTCAAGTGAAGATGTTGGAGTGTCTGTGATGATCGGATATGAGCTGAGG 2100
Db 2041 TGCAGTCAAGTGAAGATGTTGGAGTGTCTGTGATGATCGGATATGAGCTGAGG 2100
Qy 2101 CTTGTGGTGAAGCAGCAGTGTGCTGGAAAGTTGAGGTGAATGCCAGGGTCCCGTG 2160
Db 2101 CTTGTGGTGAAGCAGCAGTGTGCTGGAAAGTTGAGGTGAATGCCAGGGTCCCGTG 2160
Qy 2161 GGAATTCCTGTCTTAATGGCTGGGAATGAACATTTGCTGAAGTTGTTGCAGGCAACTT 2220
Db 2161 GGAATTCCTGTCTTAATGGCTGGGAATGAACATTTGCTGAAGTTGTTGCAGGCAACTT 2220
Qy 2221 GAATGTGGGCTGCAATCAGGCTCTCCAGAGGCTCATTTTCACAGAAAGAACATTAAC 2280
Db 2221 GAATGTGGGCTGCAATCAGGCTCTCCAGAGGCTCATTTTCACAGAAAGAACATTAAC 2280
Qy 2281 ATCTTAATGCTCAATTCGGCTGCACGTGGAGGGAAGCCTCTCTCGGATGTATAGGA 2340
Db 2281 ATCTTAATGCTCAATTCGGCTGCACGTGGAGGGAAGCCTCTCTCGGATGTATAGGA 2340
Qy 2341 TGGAGTGAACACAGCTGCTGTCAATTAATATGAAGCAAGTTTTCATCTGCTCAGCC 2400
Db 2341 TGGAGTGAACACAGCTGCTGTCAATTAATATGAAGCAAGTTTTCATCTGCTCAGCC 2400
Qy 2401 CACAGGACCCAGCTGGTGGAGCTGATATGCCCTCTCTGAGCAGTGTGAAAGTGA 2460
Db 2401 CACAGGACCCAGCTGGTGGAGCTGATATGCCCTCTCTGAGCAGTGTGAAAGTGA 2460
Qy 2461 CATCAGACATGCGCTCTGTGTGATCTGATTTCTCTCTCATGCTGCCAATGTG 2520
Db 2461 CATCAGACATGCGCTCTGTGTGATCTGATTTCTCTCTCATGCTGCCAATGTG 2520
Qy 2521 CTGTGCAGAGAAATTAATTTGGAGATGCCATATCTCTTTCTGTGGAGATCACTTTGGA 2580
Db 2521 CTGTGCAGAGAAATTAATTTGGAGATGCCATATCTCTTTCTGTGGAGATCACTTTGGA 2580
Qy 2581 AAAGGAAATGCTAACTTTGGGCGGAAAGTTCCAGTGTGAAGGAGTGAACCTCACCTT 2640
Db 2581 AAAGGAAATGCTAACTTTGGGCGGAAAGTTCCAGTGTGAAGGAGTGAACCTCACCTT 2640
Qy 2641 GCATATCCCCCATTTTCAACATCCGGAAGACACTTGTATCCACAGCAGAGAGATTGGA 2700
Db 2641 GCATATCCCCCATTTTCAACATCCGGAAGACACTTGTATCCACAGCAGAGAGATTGGA 2700
Qy 2701 GTTGTCTCTCCGATATACAGATGCTCCGACTTGTGAATGCAATCCAGTGTGACGGG 2760
Db 2701 GTTGTCTCTCCGATATACAGATGCTCCGACTTGTGAATGCAATCCAGTGTGACGGG 2760
Qy 2761 CAAGTGGAGATCAAGCTGCTTGGACACTGGGCTCACTGTGTGACACCCACTGGGACCCA 2820
Db 2761 CAAGTGGAGATCAAGCTGCTTGGACACTGGGCTCACTGTGTGACACCCACTGGGACCCA 2820
Qy 2821 GAAGATGCCCTGTCTATGACAGACAGCTCAGCTGTGGGACTGCTCTCAACCCACAGGA 2880
Db 2821 GAAGATGCCCTGTCTATGACAGACAGCTCAGCTGTGGGACTGCTCTCAACCCACAGGA 2880

Qy 2881 GGAAATATATTGGAGAAAGAGTTCGTGTGGGACACAGGTTTCATTGCTTAGGG 2940
Db 2881 GGAAATATATTGGAGAAAGAGTTCGTGTGGGACACAGGTTTCATTGCTTAGGG 2940
Qy 2941 AATGAGTCACCTTCCTGGATAACTGTCAAAATGACAGTTCCTGGAGCACTCCCTGTATCAT 3000
Db 2941 AATGAGTCACCTTCCTGGATAACTGTCAAAATGACAGTTCCTGGAGCACTCCCTGTATCAT 3000
Qy 3001 GGAATACTGTCTGTGATCTGCACAGAAAGCCCTGACCCAGCCACTGTTTCCATGCCCTC 3060
Db 3001 GGAATACTGTCTGTGATCTGCACAGAAAGCCCTGACCCAGCCACTGTTTCCATGCCCTC 3060
Qy 3061 GCAATATGATCTGACCCCATATTTGTCTCAGTTCACAGAGGCGAGTGTTCATCTGCTTA 3120
Db 3061 GCAATATGATCTGACCCCATATTTGTCTCAGTTCACAGAGGCGAGTGTTCATCTGCTTA 3120
Qy 3121 GAGGACAAACGGCTCCGCTAGTGGATGGGACAGCCGCTGTGCCGGGAGAGTAGATC 3180
Db 3121 GAGGACAAACGGCTCCGCTAGTGGATGGGACAGCCGCTGTGCCGGGAGAGTAGATC 3180
Qy 3181 TATCAGCAGCGCTTCTGGGCGACCATCTGTGATGACGCTGGGACCTGAGCGATCCCATC 3240
Db 3181 TATCAGCAGCGCTTCTGGGCGACCATCTGTGATGACGCTGGGACCTGAGCGATCCCATC 3240
Qy 3241 GTGTGTCTCAAAAGCTGGGCTGTGGAGTGGCTTTCAATGCCACGCTCTCTGCTCACATTT 3300
Db 3241 GTGTGTCTCAAAAGCTGGGCTGTGGAGTGGCTTTCAATGCCACGCTCTCTGCTCACATTT 3300
Qy 3301 GGGAGGGGTTCAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGAACGAGTCCAC 3360
Db 3301 GGGAGGGGTTCAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGAACGAGTCCAC 3360
Qy 3361 TTGTGGCAGTCCCTTCCCGGGCTGGGCGACGACGACTGCAGGCAACAAGGAGGACCA 3420
Db 3361 TTGTGGCAGTCCCTTCCCGGGCTGGGCGACGACGACTGCAGGCAACAAGGAGGACCA 3420
Qy 3421 GGGGTCTCTGCTCAGAAATTCACAGCTTTGAGGCTTACAGTGAACCTGAAACAGAGAGC 3480
Db 3421 GGGGTCTCTGCTCAGAAATTCACAGCTTTGAGGCTTACAGTGAACCTGAAACAGAGAGC 3480
Qy 3481 TGTGTGGGAGATTGGAAGTCTTCTATACGGGACCTGGGCGAGCTGGCAGGAGGAAC 3540
Db 3481 TGTGTGGGAGATTGGAAGTCTTCTATACGGGACCTGGGCGAGCTGGCAGGAGGAAC 3540
Qy 3541 ATCACCACAGCATAGCAGGATCTGTGTGACAGGCTGGGCTGTGGGAGAAATGGAGTT 3600
Db 3541 ATCACCACAGCATAGCAGGATCTGTGTGACAGGCTGGGCTGTGGGAGAAATGGAGTT 3600
Qy 3601 CTCAGCTCGCCCTTTATCTAAGACAGGCTCTGTTTCTATGTGGGTGGATGACATTCAG 3660
Db 3601 CTCAGCTCGCCCTTTATCTAAGACAGGCTCTGTTTCTATGTGGGTGGATGACATTCAG 3660
Qy 3661 TGTCTAAACGCATATCTCCATATGGCAGTCCCTGTCTGCCCATGGGAGGAGAAATC 3720
Db 3661 TGTCTAAACGCATATCTCCATATGGCAGTCCCTGTCTGCCCATGGGAGGAGAAATC 3720
Qy 3721 TCCAGCCAGCAGAGAGACCTTGAATCATGTGAAGATAGAAATAGAGTGGCTGGAGGA 3780
Db 3721 TCCAGCCAGCAGAGAGACCTTGAATCATGTGAAGATAGAAATAGAGTGGCTGGAGGA 3780
Qy 3781 GACACCGAGTGTCTGGGAGAGTGGAGATCTGGCACGAGGCTCTGGGGCACAGTGTGT 3840
Db 3781 GACACCGAGTGTCTGGGAGAGTGGAGATCTGGCACGAGGCTCTGGGGCACAGTGTGT 3840
Qy 3841 GATGACTCCTGGGACCTGGCGAGGCGGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3900
Db 3841 GATGACTCCTGGGACCTGGCGAGGCGGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3900
Qy 3901 GCTCTGGCTGCCCTGAGGAGCGCTTCGTTTGGCCAGGGAACCTGGAACCATCTGGTTGGAT 3960
Db 3901 GCTCTGGCTGCCCTGAGGAGCGCTTCGTTTGGCCAGGGAACCTGGAACCATCTGGTTGGAT 3960

```
QY 3961 GACATCGGTGCAAGAGAAATGAGTCAATTTCTATGGGACTGTACGCCCAAAACCTGGGGA 4020
Db 3961 GACATCGGTGCAAGAGAAATGAGTCAATTTCTATGGGACTGTACGCCCAAAACCTGGGGA 4020
QY 4021 CAGAGTCACTGTGGACACAGGAAGATGCTGGCGTCAGTGTCTGGACAGTCGCTGAAA 4080
Db 4021 CAGAGTCACTGTGGACACAGGAAGATGCTGGCGTCAGTGTCTGGACAGTCGCTGAAA 4080
QY 4081 TCACCTGAATGCCCTCCCTCAGTCAATTTAGCACATTTATTTATCCAGTATCTTTGGGCTCCCT 4140
Db 4081 TCACCTGAATGCCCTCCCTCAGTCAATTTAGCACATTTATTTATCCAGTATCTTTGGGCTCCCT 4140
QY 4141 CTCCTGGTCTGTTTATTTCTATTTCTACGTCGTGGCGAGTTCAGAAACAAAACATCTG 4200
Db 4141 CTCCTGGTCTGTTTATTTCTATTTCTACGTCGTGGCGAGTTCAGAAACAAAACATCTG 4200
QY 4201 CCCTCAGAGTTTCAACCAAGAGAGGGGCTCTCTCGAGGAGAAATTTATTCATGAGATG 4260
Db 4201 CCCTCAGAGTTTCAACCAAGAGAGGGGCTCTCTCGAGGAGAAATTTATTCATGAGATG 4260
QY 4261 GAGACCTGCTCAAGAGAGAGGAGCCACATGGGACAAAGACCTCAGATGACACCCCAAC 4320
Db 4261 GAGACCTGCTCAAGAGAGAGGAGCCACATGGGACAAAGACCTCAGATGACACCCCAAC 4320
QY 4321 CATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTCTTCTCGCTCTGAAGCC 4380
Db 4321 CATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTCTTCTCGCTCTGAAGCC 4380
QY 4381 ACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAG 4440
Db 4381 ACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAG 4440
QY 4441 ACAACAACTTTTAAATGAATAAGAGAGAGTCAAGTTGGCCCTATGGAAACCTTTGCCAAA 4500
Db 4441 ACAACAACTTTTAAATGAATAAGAGAGAGTCAAGTTGGCCCTATGGAAACCTTTGCCAAA 4500
QY 4501 TACATTTCTTGAAACATAGAGAGACAGCTAAATTTGATAAAGACTGGTGATAATAAAT 4560
Db 4501 TACATTTCTTGAAACATAGAGAGACAGCTAAATTTGATAAAGACTGGTGATAATAAAT 4560
QY 4561 TGAATATGATATCACTGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4620
Db 4561 TGAATATGATATCACTGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4620
QY 4621 GGGTCGAC 4628
Db 4621 GGGTCGAC 4628

RESULT 2
AAF45124
ID AAF45124 standard; cdna; 4359 BP.
XX
AC AAF45124;
XX
DT 30-MAR-2001 (first entry)
XX
DE Human TANGO 234 ORF.
XX
KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety; ss.
XX
OS Homo sapiens.
XX
PN WO200077239-A2.
XX
PD 21-DEC-2000.
XX
PF 24-MAY-2000; 2000WO-US14858.
```

```
XX 14-JUN-1999; 99US-0333159.
PR (MILL-) MILLENNIUM PHARM INC.
XX
PA McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
PI P-PSDB; AAB66037.
XX
DR WPI; 2001-032313/04.
XX
PS P-PSDB; AAB66037.
XX
TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
screening assays and diagnostic assays and for the treatment of
neurological diseases such as Alzheimer's, Parkinson's and Huntington's
disease.
XX
Claim 1; Fig 2; 359pp; English.
XX
CC The present invention relates to TANGO or INTERCEPT proteins and coding
sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
sequences are useful for the treatment of neurological disorders such as
central nervous system (CNS) disorders, CNS-related disorders, focal
brain disorders, global-diffuse cerebral disorders and other
neurological and cerebrovascular disorders. The CNS disorders include
Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
autonomic function disorders such as hypertension and sleep disorders,
neuropsychiatric disorders, psychoactive substance use disorders,
anxiety, and bipolar affective disorder.
XX
SQ Sequence 4359 BP; 1060 A; 881 C; 1303 G; 1115 T; 0 other;
Query Match 94.2%; Score 4359; DB 22; Length 4359;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 28 ATGATGCTGCTCAAAACCTCGTGGCAATATGATTTTGGAGATGCTGCTCATCAGAAC 87
Db 1 ATGATGCTGCTCAAAACCTCGTGGCAATATGATTTTGGAGATGCTGCTCATCAGAAC 60
QY 88 CTTTTCCTGCTGCTGTAACCTTCATCCTGCTGCTCAATTCCTCTTCATCAGCAGT 147
Db 61 CTTTTCCTGCTGCTGTAACCTTCATCCTGCTGCTCAATTCCTCTTCATCAGCAGT 120
QY 148 TTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 207
Db 121 TTTAATGGAACAGATTTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 180
QY 208 GTGGAGGTGAAATTCAGGACAGTGGGGACCTGTGTGATCATGATGGTGGAAACACTACT 267
Db 181 GTGGAGGTGAAATTCAGGACAGTGGGGACCTGTGTGATCATGATGGTGGAAACACTACT 240
QY 268 GCCTCAACTGCTGCTGCAAAACAGTTCGATGCTCCATTTCTTCGCAATGTTTCGTTTT 327
Db 241 GCCTCAACTGCTGCTGCAAAACAGTTCGATGCTCCATTTCTTCGCAATGTTTCGTTTT 300
QY 328 GGACAAAGCCGTGACTAGACATGGAATAATTTGCTTCATGATGTTTCTCTGTTATGAAAT 387
Db 301 GGACAAAGCCGTGACTAGACATGGAATAATTTGCTTCATGATGTTTCTCTGTTATGAAAT 360
QY 388 GAGTCACTCTTCGGAATGTCACACCGGAAATGGGAGGCCAATCACTTTATCATGGA 447
Db 361 GAGTCACTCTTCGGAATGTCACACCGGAAATGGGAGGCCAATCACTTTATCATGGA 420
QY 448 GAAGATGTTGGTGAACCTGTTATGTTGAAGCAATCTGGGTTTGGAGCTAGTGGATGGA 507
Db 421 GAAGATGTTGGTGAACCTGTTATGTTGAAGCAATCTGGGTTTGGAGCTAGTGGATGGA 480
QY 508 AACAACTCTCTTCAGGAGAGTGGAGGTGAAATTTCCAGAAAGGTGGGGACTATATGT 567
Db 481 AACAACTCTCTTCAGGAGAGTGGAGGTGAAATTTCCAGAAAGGTGGGGACTATATGT 540
QY 568 GATGATGGGTGGAACCTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 627
```

Db 541 GATGATGGGTGAACCTTCAATACCTGCTCCCTGGTGTGCAGCGCAACTAGGATGTCCATCT 600
Qy 628 TCTTTTATTTCTCTGGAGTGTGTAATAGCCCTGCTGTATTGGCCCCCAATTGGCTGAT 687
Db 601 TCTTTTATTTCTCTGGAGTGTGTAATAGCCCTGCTGTATTGGCCCCCAATTGGCTGAT 660
Qy 688 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTCAGACATCGTGGATGGGA 747
Db 661 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTCAGACATCGTGGATGGGA 720
Qy 748 AATCATGACTGAGTCACAATGAGGATGTCAATTAACCTGTTATGATAGTAGTATCTT 807
Db 721 AATCATGACTGAGTCACAATGAGGATGTCAATTAACCTGTTATGATAGTAGTATCTT 780
Qy 808 GAACCTAAGGCTGTAGTGGTAACCTGTAACCCGTGTATGGGGAGAGTAGAGCTGAATCCAA 867
Db 781 GAACCTAAGGCTGTAGTGGTAACCTGTAACCCGTGTATGGGGAGAGTAGAGCTGAATCCAA 840
Qy 868 GGAAGGTGGGGACCGTATGCCACCATTAAGTGGAAATGCTGAGCTGATGTCGTATGC 927
Db 841 GGAAGGTGGGGACCGTATGCCACCATTAAGTGGAAATGCTGAGCTGATGTCGTATGC 900
Qy 928 AAGCAGTTGGGATGGAACCGCACTTCACCTCGCTGGCTTGCCTCATTTGCAAGTCAGG 987
Db 901 AAGCAGTTGGGATGGAACCGCACTTCACCTCGCTGGCTTGCCTCATTTGCAAGTCAGG 960
Qy 988 TCTGATGTTGATGGCTGTAGTGGTGTCTCTGCTCCGCTGAATGAATCTTTCTTTGGAC 1047
Db 961 TCTGATGTTGATGGCTGTAGTGGTGTCTCTGCTCCGCTGAATGAATCTTTCTTTGGAC 1020
Qy 1048 TCCAGACATTCGGGAACCGTCAATTTTGAATGCTCTTCATCAAAAGCATGTCTGTGATC 1107
Db 1021 TCCAGACATTCGGGAACCGTCAATTTTGAATGCTCTTCATCAAAAGCATGTCTGTGATC 1080
Qy 1108 TGCTCAGATGAGCAGATTTGGAACCTGCGACTAGCAGATGGAATGAATTTGTCAGG 1167
Db 1081 TGCTCAGATGAGCAGATTTGGAACCTGCGACTAGCAGATGGAATGAATTTGTCAGG 1140
Qy 1168 AGAGTAGAGGTGAGAAATTCATGAACAGTGGTGGCAATATGTGACCAAGCTGGAAGAT 1227
Db 1141 AGAGTAGAGGTGAGAAATTCATGAACAGTGGTGGCAATATGTGACCAAGCTGGAAGAT 1200
Qy 1228 GAACAAGCCCTGTGGTTTGTGAAGCAGCTAGGATGCTCGCTTGGCAGTCTTGGCAGTCT 1287
Db 1201 GAACAAGCCCTGTGGTTTGTGAAGCAGCTAGGATGCTCGCTTGGCAGTCTTGGCAGTCT 1260
Qy 1288 CGTGCTAAACCTAGTAATGAAGCTAGACATATTTGGATTAACAGCATATCTTTGCACTGG 1347
Db 1261 CGTGCTAAACCTAGTAATGAAGCTAGACATATTTGGATTAACAGCATATCTTTGCACTGG 1320
Qy 1348 AATGAGTCAGCTCTCTGGGACTGACATATGATGAAAGCAAGCAAGCAATGCTTCCGA 1407
Db 1321 AATGAGTCAGCTCTCTGGGACTGACATATGATGAAAGCAAGCAAGCAATGCTTCCGA 1380
Qy 1408 AGATCAGATGCTGGAGTAATTTGTTCTGATAGGCAAGTCTGGACCTAAGGCTTGTCCGG 1467
Db 1381 AGATCAGATGCTGGAGTAATTTGTTCTGATAGGCAAGTCTGGACCTAAGGCTTGTCCGG 1440
Qy 1468 GCTATAGCCCTGTTATGGGAGATGGAGTGAATACCAAGGAGATGGGGGACTGTG 1527
Db 1441 GCTATAGCCCTGTTATGGGAGATGGAGTGAATACCAAGGAGATGGGGGACTGTG 1500
Qy 1528 TGTATGATCAGATGAGGACAGGAATGACGCTGTTGTGTAACAAATGGGATGTGA 1587
Db 1501 TGTATGATCAGATGAGGACAGGAATGACGCTGTTGTGTAACAAATGGGATGTGA 1560
Qy 1588 AAGCCTATGCATGTTTGGTATGACCTATTTTAAAGAGCATCAGGACCTATTTGGCTG 1647
Db 1561 AAGCCTATGCATGTTTGGTATGACCTATTTTAAAGAGCATCAGGACCTATTTGGCTG 1620
Qy 1648 GATGACGTTTCTGCAATTTGGAAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATG 1707

Db 1621 GATGACGTTTCTTGCATTTGGAAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATG 1680
Qy 1708 GGAAGCAATAATGTGTACACAGAGAGATGTGATTTGAACCTGCTCAGGTGATCAACA 1767
Db 1681 GGAAGCAATAATGTGTACACAGAGAGATGTGATTTGAACCTGCTCAGGTGATCAACA 1740
Qy 1768 TGGGSCCTGAGGCTGTGGTGGGGGCGAGCAACCGCTCTCGGGAAGACTGAGAGTGTACTTT 1827
Db 1741 TGGGSCCTGAGGCTGTGGTGGGGGCGAGCAACCGCTCTCGGGAAGACTGAGAGTGTACTTT 1800
Qy 1828 CAAGCACGCTGGGCGACAGTGTGTGATCAGCGCTGGAACAGTAAGCTGACAGCTGTGGT 1887
Db 1801 CAAGCACGCTGGGCGACAGTGTGTGATCAGCGCTGGAACAGTAAGCTGACAGCTGTGGT 1860
Qy 1888 TGTAGCCAGCTGGACTGCCCATCTTCTATCATTTGATGGTGGTGGAAACCTTCTACA 1947
Db 1861 TGTAGCCAGCTGGACTGCCCATCTTCTATCATTTGATGGTGGTGGAAACCTTCTACA 1920
Qy 1948 GGATATGAAAAAATTTGGCTCGATGATGTTTCTGTGATGGAGATGAGTCAGATCTCTGG 2007
Db 1921 GGATATGAAAAAATTTGGCTCGATGATGTTTCTGTGATGGAGATGAGTCAGATCTCTGG 1980
Qy 2008 TCATGCAGAACAGTGGGTGGGGAATTAATGACTCAGCTCAGCTCAGAGTGAAGATGTTGGAGT 2067
Db 1981 TCATGCAGAACAGTGGGTGGGGAATTAATGACTCAGCTCAGCTCAGAGTGAAGATGTTGGAGT 2040
Qy 2068 ATCTGTTCTGATGATCGATATGAGCTGAGGCTTGTGGGTGGAGCAGCAGTGTGCT 2127
Db 2041 ATCTGTTCTGATGATCGATATGAGCTGAGGCTTGTGGGTGGAGCAGCAGTGTGCT 2100
Qy 2128 GGAAGAATGAGTGAATGTTCCAGGTCCTGGGAATTCGTGTGCTTAATGGCTGGGA 2187
Db 2101 GGAAGAATGAGTGAATGTTCCAGGTCCTGGGAATTCGTGTGCTTAATGGCTGGGA 2160
Qy 2188 ATGAACATGCTGAAGTGTGTTGAGGCAACTTGAATGTGGTCTGCAATCAGGCTCTCC 2247
Db 2161 ATGAACATGCTGAAGTGTGTTGAGGCAACTTGAATGTGGTCTGCAATCAGGCTCTCC 2220
Qy 2248 AGAGAGCTCTATTTACAGAAAGAAACATTAACATCTTAAATGCGAATCTCGCTGCACT 2307
Db 2221 AGAGAGCTCTATTTACAGAAAGAAACATTAACATCTTAAATGCGAATCTCGCTGCACT 2280
Qy 2308 GAGGGGAAGCCTCTCTCTGGGATTTGATACGATGGGAGTGGAAACAGACTCGCTGTCTAT 2367
Db 2281 GAGGGGAAGCCTCTCTCTGGGATTTGATACGATGGGAGTGGAAACAGACTCGCTGTCTAT 2340
Qy 2368 TTAATATATGGAAGCAAGTGTGATCTCAGCCACAGCAGCCAGGCTGTGAGCT 2427
Db 2341 TTAATATATGGAAGCAAGTGTGATCTCAGCCACAGCAGCCAGGCTGTGAGCT 2400
Qy 2428 GATATGCCCTGCTCTGGAGCTGTTGAAAGTGAACATGAGACACATGCGCTCTCTGT 2487
Db 2401 GATATGCCCTGCTCTGGAGCTGTTGAAAGTGAACATGAGACACATGCGCTCTCTGT 2460
Qy 2488 GATCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGAGAGATTAATTTGTGGAGAT 2547
Db 2461 GATCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTGAGAGATTAATTTGTGGAGAT 2520
Qy 2548 CCCATATCTTTCTCTGGAGATCATTGGAAGGGAATGGTCTAACTTTGGGCCGAA 2607
Db 2521 CCCATATCTTTCTCTGGAGATCATTGGAAGGGAATGGTCTAACTTTGGGCCGAA 2580
Qy 2608 AAGTTCCAGTGTGAAGGAGTGAACCTCACCCTTGCATTTATGCCCATTTGTTCAACATCG 2667
Db 2581 AAGTTCCAGTGTGAAGGAGTGAACCTCACCCTTGCATTTATGCCCATTTGTTCAACATCG 2640
Qy 2668 GAAGACACTTGTATCCACAGCAGAGAGTGTGCTGTGTTCCCGATATACAGATGTC 2727
Db 2641 GAAGACACTTGTATCCACAGCAGAGAGTGTGCTGTGTTCCCGATATACAGATGTC 2700
Qy 2728 GCACCTGTAATGGCAAAATCCAGTGTGACGGCAGTGGAGATCAACCTGCTTGGACAC 2787
Db 2701 GCACCTGTAATGGCAAAATCCAGTGTGACGGCAGTGGAGATCAACCTGCTTGGACAC 2760

QY	2788	TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTCTTATCGACACAG	2847
Db	2761	TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAGATGCCCGTGTCTTATCGACACAG	2820
QY	2848	CTCAGCTGTGGACTGCTCTCTCAACACACAGAGAGAAATATATTGGAGAAAGAGTGT	2907
Db	2821	CTCAGCTGTGGACTGCTCTCTCAACACACAGAGAGAAATATATTGGAGAAAGAGTGT	2880
QY	2908	CGTGTGGGGACACAGGTTTCAATTGCTTAGGGAATGAGTCACCTCTCGGATACTGTCAA	2967
Db	2881	CGTGTGGGGACACAGGTTTCAATTGCTTAGGGAATGAGTCACCTCTCGGATACTGTCAA	2940
QY	2968	ATGACAGTTCTTTGGAGACCTCCCTGTATCCATGCGAAATACGTGCTCTGTGATCTGCACA	3027
Db	2941	ATGACAGTTCTTTGGAGACCTCCCTGTATCCATGCGAAATACGTGCTCTGTGATCTGCACA	3000
QY	3028	GGAGCCTGACCCAGCCCACTGTTTCCATGCGCTCGCAATGATCTGACCCATATTGTCT	3087
Db	3001	GGAGCCTGACCCAGCCCACTGTTTCCATGCGCTCGCAATGATCTGACCCATATTGTCT	3060
QY	3088	GCAGTTCCAGAGGGCAGTCTTGTATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGAT	3147
Db	3061	GCAGTTCCAGAGGGCAGTCTTGTATCTGCTTAGAGGACAAACGGCTCCGCCTAGTGGAT	3120
QY	3148	GGGACAGCGCTGTGCCGGGAGATAGAGATCTATCACAGCGCTCTCTGGGCAACCATC	3207
Db	3121	GGGACAGCGCTGTGCCGGGAGATAGAGATCTATCACAGCGCTCTCTGGGCAACCATC	3180
QY	3208	TGTGATCAGGCTGGGACCTGAGCATGCCACAGCTGTGTCTAAAAGCTGGCGTGTGGA	3267
Db	3181	TGTGATCAGGCTGGGACCTGAGCATGCCACAGCTGTGTCTAAAAGCTGGCGTGTGGA	3240
QY	3268	GTGGCTTCAAAGCCACGCTCTGCTCACTTTGGGAGGGGTGAGGGCCCATCTGGCGTG	3327
Db	3241	GTGGCTTCAAAGCCACGCTCTGCTCACTTTGGGAGGGGTGAGGGCCCATCTGGCGTG	3300
QY	3328	GATGACCTGAACCTGACAGGAACGGAGTCCCACTTTGTGGCAGTGCCTTCCCGGGCTGG	3387
Db	3301	GATGACCTGAACCTGACAGGAACGGAGTCCCACTTTGTGGCAGTGCCTTCCCGGGCTGG	3360
QY	3388	GGGCACACGACTGCAGGCACAGGAGGACCGAGGGTCACTGCTCAGAAATTCACAGCC	3447
Db	3361	GGGCACACGACTGCAGGCACAGGAGGACCGAGGGTCACTGCTCAGAAATTCACAGCC	3420
QY	3448	TTGAGGCTCTACAGTGAATACTGAAACAGAGAGCTGTGTGGAGATTGGAGTCTTCTAT	3507
Db	3421	TTGAGGCTCTACAGTGAATACTGAAACAGAGAGCTGTGTGGAGATTGGAGTCTTCTAT	3480
QY	3508	AACGGGACCTGGGGACGCTGGCAGGAGGAGATCATCCACAGCCATAGAGGCAATTGTG	3567
Db	3481	AACGGGACCTGGGGACGCTGGCAGGAGGAGATCATCCACAGCCATAGAGGCAATTGTG	3540
QY	3568	TGACGCAGCTGGGCTGTGGGAGAAATGAGTGTGCAGCTCGGCCCTTTATCTAAGACA	3627
Db	3541	TGACGCAGCTGGGCTGTGGGAGAAATGAGTGTGCAGCTCGGCCCTTTATCTAAGACA	3600
QY	3628	GGCTCTGGTTTCAATGTGGGTGGATGACATTCAGTGTCTCTAAACGCATATCTCCATATGG	3687
Db	3601	GGCTCTGGTTTCAATGTGGGTGGATGACATTCAGTGTCTCTAAACGCATATCTCCATATGG	3660
QY	3688	CAGTGCCTGTCTGCCCATGGGAGCGAAGATCTCCAGCCAGACAGAGACCTGGATC	3747
Db	3661	CAGTGCCTGTCTGCCCATGGGAGCGAAGATCTCCAGCCAGACAGAGACCTGGATC	3720
QY	3748	ACATGTGAAGATAGAATAAGTGTGGGAGGACACCGAGTGCCTCTGGGAGAGTGGAG	3807
Db	3721	ACATGTGAAGATAGAATAAGTGTGGGAGGACACCGAGTGCCTCTGGGAGAGTGGAG	3780
QY	3808	ATCTGGCACCGAGCTCTCTGGGACACAGTGTGTGATGACTCTCTGGGACCTGCCGAGGCG	3867
Db	3781	ATCTGGCACCGAGCTCTCTGGGACACAGTGTGTGATGACTCTCTGGGACCTGCCGAGGCG	3840

RECEIVED 3

RESUL 3
AASC01217

AAS0121/
ID AAS01217 standard: cDNA: 3670 BP.

ID
 YY
 AASU1211/

XX
AC
AAC01217.

AC
VV
AASULZ17;

XX.
DT 04-III.-2001 (first entry)

DT 04-JUL-2001 (LIST ENCLY)
VY

XX DNA encoding human secreted protein. P01.Y8.

DE
YY
DNA encoding human secreted protein, F0010:

XX human secreted protein: therapeutic: diagnostic: human: cancer: ss.

Human secreted protein; C

[illegible]

OS Homo sapiens.

[illegible]

EH	Key	Location/
113	3475	

FT CDS 113..34/5

```
FT
/*tag= a
```

FT=product=

XX
XX
C
A
U
L
O
C
E
R
C
O
C
T
I

PN WO200119856-A2.

XX
XX

PD 22-MAR-2001.

XX

PF 13-SEP-2000; 2000WO-US2510

XX

DR P-PSDB; AAU00396.
XX New POLYX polypeptide useful for treating or preventing a POLYX
PT associated disorder, e.g. cancer
XX
XX
PS Claim 9; Page 25-29; 152pp; English.
XX
XX The sequence represents the coding sequence of human secreted protein,
CC POLYX. POLYX nucleic acids, polypeptides and antibodies to POLYX can be
CC used for treating or preventing a POLYX associated disorder in a subject,
CC preferably a human. These can be used in the manufacture of a medicament
CC for treating a syndrome associated with a human disease selected from a
CC POLYX-associated disorder, where the therapeutic is a POLYX polypeptide,
CC a POLYX nucleotide or a POLYX antibody. They may also be used to screen
CC for a modulator of activity, or latency, or predisposition to a POLYX
CC associated disorder, e.g. cancer.
XX
SQ Sequence 3670 BP; 947 A; 752 C; 1073 G; 898 T; 0 other;

Query Match 74.8%; Score 3459.8; DB 22; Length 3670;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3464; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1114 GATGAGCAGATTGGAACTCGGACTAGCAGATGGAAGTACAATTGTTACGGGAGAGTA 1173
Db 200 GATGAGCAGATTGGAACTCGGACTAGCAGATGGAAGTACAATTGTTACGGGAGAGTA 259

Qy 1174 GAGGTGAGAATTCATGAACAGTGTGGACAATATGTGACCAGAACTGGAAGAATCAACAA 1233
Db 260 GAGGTGAGAATTCATGAACAGTGTGGACAATATGTGACCAGAACTGGAAGAATCAACAA 319

Qy 1234 GCCCTTGTGGTGTGTAAGCAGCTAGGATGTCGTTTCAGCGTCTTTGGCAGTCTGCTGCT 1293
Db 320 GCCCTTGTGGTGTGTAAGCAGCTAGGATGTCGTTTCAGCGTCTTTGGCAGTCTGCTGCT 379

Qy 1294 AAACCTAGTAATGAAGCTAGAGACATTTGGATAACACGCATATCTTGCACTGGGAATGAG 1353
Db 380 AAACCTAGTAATGAAGCTAGAGACATTTGGATAACACGCATATCTTGCACTGGGAATGAG 439

Qy 1354 TCAGCTCTCTGGGACTGCACATATGATGAAAGCAAGCAACATGCTTCCGGAAGATCA 1413
Db 440 TCAGCTCTCTGGGACTGCACATATGATGAAAGCAAGCAACATGCTTCCGGAAGATCA 499

Qy 1414 GATGCTGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTTAAGCCTTGTCCGGGCTCAT 1473
Db 500 GATGCTGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTTAAGCCTTGTCCGGGCTCAT 559

Qy 1474 AGCCCTCTTATGGAGATTGGAGTGAATACCAAGAGAGTGGGGACATGTGTGTCAT 1533
Db 560 AGCCCTCTTATGGAGATTGGAGTGAATACCAAGAGAGTGGGGACATGTGTGTCAT 619

Qy 1534 GACAGATGGAGCACAAAGGAATGCAGCTGTGTGTGTAACAAATTTGGGATGTGGAAGCCT 1593
Db 620 GACAGATGGAGCACAAAGGAATGCAGCTGTGTGTGTAACAAATTTGGGATGTGGAAGCCT 679

Qy 1594 ATGCATGTGTTTGGTATGACCTATTTTAAAGAGCATCAGGACCTATTGTCGTGATGAC 1653
Db 680 ATGCATGTGTTTGGTATGACCTATTTTAAAGAGCATCAGGACCTATTGTCGTGATGAC 739

Qy 1654 GTTTCCTGCATTGGAATAGTCAAAATATCTGGGACTGTGAACACAGTGGATGGGGAAG 1713
Db 740 GTTTCCTGCATTGGAATAGTCAAAATATCTGGGACTGTGAACACAGTGGATGGGGAAG 799

Qy 1714 CATAACTGTACACAGAGAGATGTGATTTAACTGCTCAGGTGATGCAACATGGGGC 1773
Db 800 CATAACTGTACACAGAGAGATGTGATTTAACTGCTCAGGTGATGCAACATGGGGC 859

Qy 1774 CTGAGGCTGGTGGCGGCGAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGA 1833
Db 860 CTGAGGCTGGTGGCGGCGAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGA 919

Qy 1834 CGGTGGGCACAGTGTGTGATGACGGCTGGAACTGAAGCTGCACTGTGCTGTGTGAGC 1893
Db 2000 TGGGGACACAGGTTTTCATTGCTTAGGGAATGAGTCACTTCTGATAACTGTCAAAATGACA 2059

Db 920 CGGTGGGCACAGTGTGTGATGACGGCTGGAACGGTAAAGCTGACGCTGTGTGTGAGC 979
Qy 1894 CAGCTGGAGTCCCATCTTCTATCATTTGSCATGGCTGCGGAAACGCTTCTACAGATAT 1953
Db 980 CAGCTGGAGTCCCATCTTCTATCATTTGSCATGGCTGCGGAAACGCTTCTACAGATAT 1039
Qy 1954 GGAATAATTTGGCTCGATGATGTTTCTCTGTGATGAGATGAGTCAAGTCTCTGGTCATGC 2013
Db 1040 GGAATAATTTGGCTCGATGATGTTTCTCTGTGATGAGATGAGTCAAGTCTCTGGTCATGC 1099
Qy 2014 AGGAACAGTGGTGGGGAATAATCACTGCAGTCAAGTGAAGATGTTGGAGTGTCTGT 2073
Db 1100 AGGAACAGTGGTGGGGAATAATCACTGCAGTCAAGTGAAGATGTTGGAGTGTCTGT 1159
Qy 2074 TCTGATGATCGGATATGAGAGTGTGGCTGGAAGCAGCAGGCTGTCTGGAATA 2133
Db 1160 TCTGATGATCGGATATGAGAGTGTGGCTGGAAGCAGCAGGCTGTCTGGAATA 1219
Qy 2134 GTTGAGGTGAATGTCACGGGTGCCGTGGGAATTTCTGTGTCTTAATGTGCTGGAATGAAC 2193
Db 1220 GTTGAGGTGAATGTCACGGGTGCCGTGGGAATTTCTGTGTCTTAATGTGCTGGAATGAAC 1279
Qy 2194 ATTGCTGAAGTTGTTGACGGCAACTTGAATGTGGTCTGCAATCAGGCTCTCCAGAGAG 2253
Db 1280 ATTGCTGAAGTTGTTGACGGCAACTTGAATGTGGTCTGCAATCAGGCTCTCCAGAGAG 1339
Qy 2254 CCTCATTTACAGAAAGAACATTACACATCTTAAATGTGCAATTTCTGGCTGCACTGGAGG 2313
Db 1340 CCTCATTTACAGAAAGAACATTACACATCTTAAATGTGCAATTTCTGGCTGCACTGGAGG 1399
Qy 2314 GAAGCTCTCTCTGGGATTTGATACGATGGGAGTGGAAACAGACACTGCGTGTCTTAAAT 2373
Db 1400 GAAGCTCTCTCTGGGATTTGATACGATGGGAGTGGAAACAGACACTGCGTGTCTTAAAT 1459
Qy 2374 ATGGAAGCAAGTTTCTGCTCAGCCACAGCCAGCCAGCCAGCTGTTGGAGCTGATATG 2433
Db 1460 ATGGAAGCAAGTTTCTGCTCAGCCACAGCCAGCCAGCCAGCTGTTGGAGCTGATATG 1519
Qy 2434 CCCTCTCTGAGAGTGTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2493
Db 1520 CCCTCTCTGAGAGTGTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1579
Qy 2494 GATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2553
Db 1580 GATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1639
Qy 2554 TCTCTTTCTGCTGGAGATCCTTTGGAAGGAAATGCTTAACTTGGCCGCAAGATTC 2613
Db 1640 TCTCTTTCTGCTGGAGATCCTTTGGAAGGAAATGCTTAACTTGGCCGCAAGATTC 1699
Qy 2614 CAGTGTGAAGGAGTGAACACTCACCCTTGCATTTATGCCCATTTGTTCAACATCCGGAAGAC 2673
Db 1700 CAGTGTGAAGGAGTGAACACTCACCCTTGCATTTATGCCCATTTGTTCAACATCCGGAAGAC 1759
Qy 2674 ACTTGTATCCACAGCAGAGAAGTTGAGTGTCTGTTCCCGATATACAGATGTCGACTT 2733
Db 1760 ACTTGTATCCACAGCAGAGAAGTTGAGTGTCTGTTCCCGATATACAGATGTCGACTT 1819
Qy 2734 GTGAATGGCAATCCAGTGTGACGGCAAGTGGAGATCAAGCTGTTGGACACTGGGGC 2793
Db 1820 GTGAATGGCAATCCAGTGTGACGGCAAGTGGAGATCAAGCTGTTGGACACTGGGGC 1879
Qy 2794 TCACATGTGTGACACCCACTGCGGACCCAGAGATGCCGTGTTCTATGACAGACAGCTCAGC 2853
Db 1880 TCACATGTGTGACACCCACTGCGGACCCAGAGATGCCGTGTTCTATGACAGACAGCTCAGC 1939
Qy 2854 TGTGGAGTGTCTCTCAACCCAGAGGAAATATATTGGAGAAAGAGTGTCTGTGTG 2913
Db 1940 TGTGGAGTGTCTCTCAACCCAGAGGAAATATATTGGAGAAAGAGTGTCTGTGTG 1999
Qy 2914 TGGGGACACAGGTTTCAATTGCTTAGGGAATGAGTCACTTCTCGATAACTGTCAAAATGACA 2973
Db 2000 TGGGGACACAGGTTTCAATTGCTTAGGGAATGAGTCACTTCTCGATAACTGTCAAAATGACA 2059

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F, Xu C;
PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
PI Drmanac R;
XX WPI: 2001-589862/66.
DR P-PSDB: AAU27709.
XX Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis, treatment of
PT cancer, neurological, inflammatory disorders and for use in arrays for
XX detection
PS Claim 1; SEQ ID No 34; 153pp; English.
XX Sequences AAS44576-AAS44919 represent full-length polynucleotides and
CC contig polynucleotides encoding polypeptides of the invention. The DNA
CC and protein sequences are useful for the treatment, diagnosis and
CC prevention of various types of disorder in a mammalian subject such as a
CC human, dog, monkey, mouse, hamster or rat. The disorders include cancers
CC such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
CC as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
CC bowel disease. The sequences exhibit activity relating to angiogenesis,
CC cell proliferation, cell differentiation, stem cell growth factor,
CC activin or inhibin. Therefore, they can be used to manipulate stem cells
CC in culture to give rise to neuroepithelial cells that can be used to
CC augment or replace cells damaged by illness, accidental damage or genetic
CC disorders. The sequences may also be used for regeneration of bone,
CC cartilage, tendons and ligaments and in tissue repair and burn healing.
CC Note: Some sequences for this patent did not form part of the printed
CC specification, but were obtained in electronic-format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2800 BP; 697 A; 607 C; 813 G; 683 T; 0 other;

Query Match 59.7%; Score 2761.2; DB:22; Length 2800;
Best Local Similarity 99.7%; Pred. No.0;
Matches 2766; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1821 GTACTTTCAAGGACGGTGGGCGACAGTGTGTGATCAGCGCTGGAACAGTAAAGCTGCAGC 1880
DB 27 GTGACTTTTCAGGACCGTGGGCGACAGCGTGTGATCAGCGCTGGAACAGTAAAGCTGCAGC 86

QY 1881 TGTGTGTGTAGCCAGCTGGAGTGGCCCATCTTCTATCATTTGGCATGGGTCTGGGAAACGC 1940
DB 87 TGTGTGTGTAGCCAGCTGGAGTGGCCCATCTTCTATCATTTGGCATGGGTCTGGGAAACGC 146

QY 1941 TTCTACAGATATGAAAAATTTGGCTCGATGATGTTTCTCTGTGATGGAGATGAGTCAGA 2000
DB 147 TTCTACAGATATGAAAAATTTGGCTCGATGATGTTTCTCTGTGATGGAGATGAGTCAGA 206

QY 2001 TCTCTGGTTCATCGAGAACAGTGGGTGGGAAATTAATGACTCAGTCAAGTGAAGATGT 2060
DB 207 TCTCTGGTTCATCGAGAACAGTGGGTGGGAAATTAATGACTCAGTCAAGTGAAGATGT 266

QY 2061 TGGAGTGTATCTGTTCTGATGATCGGATATGGAGCTGAGGCTTGTGGGTGGGAAACAGCAG 2120
DB 267 TGGAGTGTATCTGTTCTGATGATCGGATATGGAGCTGAGGCTTGTGGGTGGGAAACAGCAG 326

QY 2121 GTGTGCTGGAAAGTTGAGTGAATGTCAGGGTGGCGTGGGAATTCGTGTGCTTAATGG 2180
DB 327 GTGTGCTGGAAAGTTGAGTGAATGTCAGGGTGGCGTGGGAATTCGTGTGCTTAATGG 386

QY 2181 CTGGGGAATGAACATTGCTGAAGTTGTTTGGCAGGCAACTTGAATGGGCTTGCATCAG 2240
DB 387 CTGGGGAATGAACATTGCTGAAGTTGTTTGGCAGGCAACTTGAATGGGCTTGCATCAG 446

QY 2241 GGTCTCCAGAGAGCCTCAATTTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCCTGG 2300

DB 447 GGTCTCCAGAGAGCGCTCAATTTTCACAGAAAGAACATTACACATCTTAATGTCGAATTCCTGG 506
QY 2301 CTGCACCTGGAGGGAGCGCTCTCTCTGGATTTGATACGATGGGAGTGAACACACACTGC 2360
DB 507 CTGCACCTGGAGGGAGCGCTCTCTCTGGATTTGATACGATGGGAGTGAACACACACTGC 566
QY 2361 GTGTCAATTAATATGGAAGCAAGTTTGTATCTGCTCAGCCACAGCGAGCCAGCTGGT 2420
DB 567 GTGTCAATTAATATGGAAGCAAGTTTGTATCTGCTCAGCCACAGCGAGCCAGCTGGT 626
QY 2421 TGGAGCTGATATGCCCTCTCTGACAGCTGTGGAAGTGAACATGACAGACATGCGGCTC 2480
DB 627 TGGAGCTGATATGCCCTCTCTGACAGCTGTGGAAGTGAACATGACAGACATGCGGCTC 686
QY 2481 TGTCTGTCTATCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTCAGAGAAATTAATTTG 2540
DB 687 TGTCTGTCTATCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTCAGAGAAATTAATTTG 746
QY 2541 TGGAGATGCCATATCTCTTTCTGTGGGAGATCACCTTTTGGAAAAGGAATGGTCTAACTTG 2600
DB 747 TGGAGATGCCATATCTCTTTCTGTGGGAGATCACCTTTTGGAAAAGGAATGGTCTAACTTG 806
QY 2601 GCGCGAAAAGTTCCAGTGTGAAGGAGTGAACACTCACCTTGCATTTATGCCCCATTTGTCA 2660
DB 807 GCGCGAAAAGTTCCAGTGTGAAGGAGTGAACACTCACCTTGCATTTATGCCCCATTTGTCA 866
QY 2661 ACATCCGGAAGACACTGTGTATCCACAGCAGAGAAGTGGAGTGTGCTGTTCCCGATATAC 2720
DB 867 ACATCCGGAAGACACTGTGTATCCACAGCAGAGAAGTGGAGTGTGCTGTTCCCGATATAC 926
QY 2721 AGATCTCCGACTTGTGAATGGCAATCCAGTGTGACGGGCAAGTGGAGATCAACGTGCT 2780
DB 927 AGATCTCCGACTTGTGAATGGCAATCCAGTGTGACGGGCAAGTGGAGATCAACGTGCT 986
QY 2781 TGGACACTGGGCTCCTGTGTGACACCCACTGGGACCCAGAGATGCCGTGTTCTATG 2840
DB 987 TGGACACTGGGCTCCTGTGTGACACCCACTGGGACCCAGAGATGCCGTGTTCTATG 1046
QY 2841 CAGACAGCTCAGCTGTGGGACTGCTCTCAACACAGAGGAGAAATATATTGGAGAAG 2900
DB 1047 CAGACAGCTCAGCTGTGGGACTGCTCTCAACACAGAGGAGAAATATATTGGAGAAG 1106
QY 2901 AAGTGTCTGTGTGGGACACAGCTTTTCATTTGCTTAGGAAATGAGTCACTTCTCGATAA 2960
DB 1107 AAGTGTCTGTGTGGGACACAGCTTTTCATTTGCTTAGGAAATGAGTCACTTCTCGATAA 1166
QY 2961 CTGTCAATGACAGTTCCTTGGAGACCTCCCTGTATCCATGGAATACTGTCTCTGTAT 3020
DB 1167 CTGTCAATGACAGTTCCTTGGAGACCTCCCTGTATCCATGGAATACTGTCTCTGTAT 1226
QY 3021 CTGCACAGAACCTTGACCCAGCCACTGTTTCCATGCCCTCGCAATGTATCTGACCCATA 3080
DB 1227 CTGCACAGAACCTTGACCCAGCCACTGTTTCCATGCCCTCGCAATGTATCTGACCCATA 1286
QY 3081 TTTGTCTGCAGTTCAGAGGGCAGTCTTGTATCTGCTTTAGAGGACAAACGGCTCCGCCT 3140
DB 1287 TTTGTCTGCAGTTCAGAGGGCAGTCTTGTATCTGCTTTAGAGGACAAACGGCTCCGCCT 1346
QY 3141 ACTGATGGGACAGCGCTGTCCGGGAGAGTACAGATCTATCAGCAGCGGCTTCTGGGG 3200
DB 1347 ACTGATGGGACAGCGCTGTCCGGGAGAGTACAGATCTATCAGCAGCGGCTTCTGGGG 1406
QY 3201 CACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCACGTGGTGTGTGTCAAAAGCTGGG 3260
DB 1407 CACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCACGTGGTGTGTGTCAAAAGCTGGG 1466
QY 3261 CTGTGGAGTGGCTTCAATGCCACAGCTCTCTGCTCACTTTTGGGGAGGGGTGAGGGCCCAT 3320
DB 1467 CTGTGGAGTGGCTTCAATGCCACAGCTCTCTGCTCACTTTTGGGGAGGGGTGAGGGCCCAT 1526
QY 3321 CTGGCTGGATGACCTGAACATGCACAGGAGTCCCACTTGTGGCAGTGCCTTCCCG 3380

Db	1527	CTGGCTGGATGACCTGAACTGCACAGAAATGGAGTCCACATTGTGGCGATGCCCCCTCCCG	1589
Qy	3381	CGGCTGGGGGACGACGACTGCAGGCACAAGGAGGACGAGGGTCTATCTGCTCAGAAATT	3440
Db	1587	CGGCTGGGGGACGACGACTGCAGGCACAAGGAGGACGAGGGTCTATCTGCTCAGAAATT	1646
Qy	3441	CACAGCCTTGAAGGCTCTACAGTGAACCTGAACAGAGAGAGCTGTGCTGGGAGATTGGAAGT	3500
Db	1647	CACAGCCTTGAAGGCTCTACAGTGAACCTGAACAGAGAGAGCTGTGCTGGGAGATTGGAAGT	1706
Qy	3501	CTTCTATAAGGGGACCTGGGGCAGCGTGGCAGGAGGAACATCACCACAGCCATAGCAGG	3560
Db	1707	CTTCTATAAGGGGACCTGGGGCAGCGTGGCAGGAGGAACATCACCACAGCCATAGCAGG	1766
Qy	3561	CATTGTGTGCAGGCAGCTGGGCTGTGGGAGAAATGGAGTGTGCAGCCTCGCCCTTTATC	3620
Db	1767	CATTGTGTGCAGGCAGCTGGGCTGTGGGAGAAATGGAGTGTGCAGCCTCGCCCTTTATC	1826
Qy	3621	TAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTTCAGTGTCTTAAAGCATATCTC	3680
Db	1827	TAAGACAGGCTCTGGTTTCATGTGGGTGGATGACATTTCAGTGTCTTAAAGCATATCTC	1886
Qy	3681	CATATGSCAGTGCCTGTCTGCCCATGGGAGGAGAAATCTCCAGGCCACGACAGAGAC	3740
Db	1887	CATATGSCAGTGCCTGTCTGCCCATGGGAGGAGAAATCTCCAGGCCACGACAGAGAC	1946
Qy	3741	CTGGATCACATGTGAAGATAGAATAAGAGTGGCTGGAGGACACCGAGTGCCTCGGGAG	3800
Db	1947	CTGGATCACATGTGAAGATAGAATAAGAGTGGCTGGAGGACACCGAGTGCCTCGGGAG	2006
Qy	3801	AGTGGAGATCTGSCACGACAGGCTCTGGGGCACAGTGTGATGACTCCTGGGACCTTGGC	3860
Db	2007	AGTGGAGATCTGSCACGACAGGCTCTGGGGCACAGTGTGATGACTCCTGGGACCTTGGC	2066
Qy	3861	CGAGGCGAAGTGGTGTGTACAGACTGGCTGTGGCTCTGCTCTGGCTGCCTTGAGGGA	3920
Db	2067	CGAGGCGAAGTGGTGTGTACAGACTGGCTGTGGCTCTGCTCTGGCTGCCTTGAGGGA	2126
Qy	3921	CGCTTCCTTTTGGCCAGGAACTGGAACCATCTGGTTGGATGACATGCGGTCAAAAGGAAA	3980
Db	2127	CGCTTCCTTTTGGCCAGGAACTGGAACCATCTGGTTGGATGACATGCGGTCAAAAGGAAA	2186
Qy	3981	TGAGTCAATFTCTATGGACTGTACGCCAAACCCCTGGGGACAGAGTGACTGTGACACAA	4040
Db	2187	TGAGTCAATFTCTATGGACTGTACGCCAAACCCCTGGGGACAGAGTGACTGTGACACAA	2246
Qy	4041	GGAAGATGCTGGCGTGAAGTGTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGG	4100
Db	2247	GGAAGATGCTGGCGTGAAGTGTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGG	2306
Qy	4101	TCATTTAGCACTTATTTATPCCAGTATCTTTGGGCTCCTTCTCCTGGTCTGTGTTTATTCT	4160
Db	2307	TCATTTAGCACTTATTTATCCAGTATCTTTGGGCTCCTTCTCCTGGTCTGTGTTTATTCT	2366
Qy	4161	ATTTCTCACGTGGTGGCGAGTTGCAAAACAAAACATCTGCCCTCAGAGTTTCAACCG	4220
Db	2367	ATTTCTCACGTGGTGGCGAGTTGCAAAACAAAACATCTGCCCTCAGAGTTTCAACCG	2426
Qy	4221	AAGGAGGGGTTCTCTCGAGGAGAAATTTATTCATGAGATGGAGACCTGCCCTCAGAGAGA	4280
Db	2427	AAGGAGGGGTTCTCTCGAGGAGAAATTTATTCATGAGATGGAGACCTGCCCTCAGAGAGA	2486
Qy	4281	GGACCCACATGGGACAAGAACCTCAGATGACACACCCCAACCATGGTTGTGAAGATGCTAG	4340
Db	2487	GGACCCACATGGGACAAGAACCTCAGATGACACACCCCAACCATGGTTGTGAAGATGCTAG	2546
Qy	4341	CGACACATCGCTGTGGAGTTCTTCTCGCTCTGAAGCCACAAAATGACTTTAGACTTC	4400
Db	2547	CGACACATCGCTGTGGAGTTCTTCTCGCTCTGAAGCCACAAAATGACTTTAGACTTC	2606
Qy	4401	CAGGGCTACCAATCAACCTCTAAATATCTTTGAAGGAGACACAACTTTTAATGAAT	4460
Db	2607	CAGGGCTACCAATCAACCTCTAAATATCTTTGAAGGAGACACAACTTTTAATGAAT	2666

Qy	4461	AAGAGAGAGCTCAAGTGGCCCTATGGAACACTTGTCAAATAACAATTCTTTGAACAATAG	4520
Db	2667	AAAGAGAGAGCTCAAGTGGCCCTATGGAACACTTGTCAAATAACAATTCTTTGAACAATAG	2726
Qy	4521	GAGAACAGCTAAATTGCATAAAGACTGGTGATTAATAAAAAATTCAAATTATGTATATCAGTGT	4580
Db	2727	GAGAACAGCTAAATTGCATAAAGACTGGTGATTAATAAAAAATTCAAATTATGTATATCAGTGT	2786
Qy	4581	TAAAAAAAAAAAAA 4594	
Db	2787	TAAAAAAAAAAAAA 2800	
RESULT 5			
ID	AAI58649	standard; cDNA; 3716 BP.	
XX	AAI58649;		
XT	22-OCT-2001	(first entry)	
XX	Human polynucleotide SEQ ID NO 852.		
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	anethropic lateral sclerosis; Shy-Drager Syndrome; chemotactic;		
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;		
XX	leukaemia; ss.		
OS	Homo sapiens.		
XX	WO200153312-A1.		
PD	26-JUL-2001.		
PF	26-DEC-2000; 2000WO-US34263.		
PR	21-JAN-2000; 2000US-0488725.		
PR	25-APR-2000; 2000US-0552317.		
PR	09-JUL-2000; 2000US-0598042.		
PR	19-JUL-2000; 2000US-0620312.		
PR	03-AUG-2000; 2000US-0653450.		
PR	14-SEP-2000; 2000US-0662191.		
PR	19-OCT-2000; 2000US-0693036.		
XX	29-NOV-2000; 2000US-0727344.		
PA	(HYSE-) HYSEQ INC.		
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;		
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;		
DR	WPI; 2001-442253/47.		
DR	P-PDSB; AAM39493.		
XX	Novel nucleic acids and polypeptides, useful for treating disorders		
PT	such as central nervous system injuries -		
XX	Claim 1; SEQ ID NO 852; 10078pp; English.		

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAAM42213) with neurotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

Sequence 3716 BP; 993 A; 730 C; 1060 G; 933 T; 0 other;

```
Query Match      31.3%; Score 1448.2; DB 22; Length 3716;
Best Local Similarity 65.4%; Pred. NO. 0;
Matches 2179; Conservative 0; Mismatches 1128; Indels 24; Gaps 3;
```

Qy	1002	GC	TTGATGGTGTCCTCTGCTCCGGTAATGAATCTTTTCTTTGGCAGTGCAGACATTC	CCGG	1061
Db	107	GCT	ACTTTGAAGACTCTGGATCTGCTGACTTCAAGAAACATTTTGTCAACTT	TGAGTCCCTT	166
Qy	1062	AAC	CGTCAATTTTGACTGTCCTTCATCAAAACGATGTGCTGTGATCTGCTCAGATG	AGC	1121
Db	167	CAC	CAATTACTTGGTCTTACTTCTCAGTGCCTTGTTTTGTCCACCACTTCTCTT	TGGAGGAAAC	226
Qy	1122	AG	ATTTGGAACTCGCACTAGCAGTGGAAAGTAACAAATTTGTCAGGAGAGAT	TAGAGTGGAG	1181
Db	227	AG	ACAAGGAGCTGAGCTAGTGGATGGTCAAAACAAGTGTACGCGGAGAGT	TGGAAGTGA	286
Qy	1182	AAT	TCAATGAACAGTGGTGGACAATATGTGACCAAGCTGGGAAGTAATGAACAAG	CCCTTGT	1241
Db	287	AGT	CCAGGAGGAGTGGGAAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGG	TCTCTGT	346
Qy	1242	GGT	TTTGAAGCAGCTAGGATGTCGGTTTCAAGCTCTTTGGCAGTCTGTCGTCTAA	ACCTAG	1301
Db	347	GAT	TTTGTAAACCACTGGGATGTCCAACTGCTATCAAAAGCCCTGGATGGGTAA	TTCAC	406
Qy	1302	TA	ATGAAGCTTAGACATATTTGGATTAACAGCATATCTTTGCACCTGGGAAT	GAGTCAGCTCT	1361
Db	407	TGC	AGGTTCTTGGACGCAATTTGGATGGATCATGTTTCTTGTCTGGGAAT	TGATCAGCTCT	466
Qy	1362	CTG	GAGCTGCACATATGATGGAAGAAACAAGC---GAACATGCTTCCGAAGAT	CAGATGC	1418
Db	467	TTG	GGATTTGCAACATGATGGATGGGAAAGCATAGTAACCTGTA	CTACCAACAAGATGC	526
Qy	1419	TG	GATTAATTTGTTCGATTAAGCAGATCTGACCTAAGGCTTGTCTGGGCTCAT	AGCCC	1478
Db	527	TG	GATGACCTGCTCAGATGGATTCAAATTTGGAAATGAGGCTGACGGTGG	AGGGAATAT	586
Qy	1479	CTG	TTATGGGAGATTTGGAGGTCAAAATACCAAGGAGAGTGGGGACTGTGCTCAT	AGAC	1538
Db	587	GTG	TTCTTGGNAGATAGATCAAAATCCAAGGAGGTTGGGACAGAGTGTGATGA	TAA	646
Qy	1539	ATG	GACACAGAAGATGCAAGCTGTTGTGTGAACAATTTGGGATGTGGAAGCCTAT	GCA	1598
Db	647	CTT	CAACATAGATCATGATCTGTCAATTTGTAGACAACCTTCAATGTGGAAGT	TGCTGCAG	706
Qy	1599	TGT	TTTGTGATGACCTATTTTAAAGAACCATCAGGACCTATTTTGGCTGGAT	GACGTTTC	1658
Db	707	TTT	CTCTGTTTCACTCAATTTTGGAGAAAGGCTCTTGACCAATCTGGTTT	GATGATCTTAT	766
Qy	1659	TTG	CAATGGAATGAGTCAAAATCTCGGACTGTGAACACAGTGGATGGGAAGCA	TAA	1718
Db	767	ATG	CAAGGAAATGAGTCAAGCTCTCTGNACTGCAACATCAAGGATGGGAAGCA	TAA	826
Qy	1719	TTG	TGTACACAGAGGATGTGATTTGAACCTGCTCAGGTGATCAACATGGG	CCCTGAG	1778
Db	827	CTG	TGATCATGCTGAGGATGCTGGAGTGATTTTGCTCAAGGAGCAGATCTG	AGCCCTGAG	886
Qy	1779	GCT	TGTTGGCGCAGCAACCGTCTCGGGAAGACTGAGGTGTACTTTCAAGGAC	GGTG	1838
Db	887	ACT	GGTAGATGGAGTCACTCAATTTGTTCAGGAGATTTAGAAGT	TGAGATTC	946
Qy	1839	GGC	CACAGTGTGTGATGACCGCTGGAACAGTAAGCTGCACGCTGTGTTGT	TAGCCAGCT	1998
Db	947	GGG	ACATATGTGATGACCGCTGGACAGTTTACGATGCTGCTGTGGCATC	CAAGCAACT	1006
Qy	1899	GG	ACTGCCCATCTTCTATCATTTGGCATGGGCTTGGAAACGCTTCTACAG	GATATGGAA	1958

Qy 1002 GCTTGATGGTCTCTCCGCTAATGAATCTTTCTTTGGGACTGCAGACATTCGGG 1061
Db 107 GCTACTTGAAGACTCTGATCTGCTGACTTCAGAAGACATTTTGTCAACCTGAGTCCCTT 166
Qy 1062 AACCGTCAATTTTGACTCTCTTCATCAAAAGCATGTGTCTGTGATCTGCTCAGATGGAGC 1121
Db 167 CACCACTACTGTGCTCTACTTCTCAGTGCCTGTTTGTCAACCACTCTCTTTGGAGAAC 226
Qy 1122 AGATTTGGAAGCTAGCAGTACAGATGAAGTAACAATTTTTCAGGGAGAGTAGAGGTGAG 1181
Db 227 AGACAAGGAGCTAGGTAGTGTGATGTGAACAAGGTGAGCGGAGAGTGAAGTGA 286
Qy 1182 AATTCAATGAACAGTGGTGGCAATATGTGACCAGAATCTGGAAGAATGAACAAGCCCTGTG 1241
Db 287 AGTCCAGGAGGTGGGACGGTGTGTATATAGTGGTGGAGCATGGAAGCGGTCTCTGT 346
Qy 1242 GGTTTGTAAGAGCTAGAGTGTCCGTTGAGGCTCTTTGGCAGTCTGTGCTGTAACCTAG 1301
Db 347 GATTTGTAACCAAGCTGGGATGTCCAACCTGCTATCAAAAGCCCTGGATGGGCTAATTCAG 406
Qy 1302 TAATGAAGCTAGACACATTTGGATAAACAGCATATCTTGGCACTGGGAATGAGTCAGCTCT 1361
Db 407 TGCAGGTTCTGGAGCGATTTGGATGGATCATGTTTCTTGTCTGGGAATGAGTCAGCTCT 466
Qy 1362 CTGGGACTGCACATATGATGGAAAGCAAAAGC --- GAACATGCTTCCGAAGATCAGATGC 1418
Db 467 TTGGATTTGCAACATGATGGATGGGNAAGCATAGTAACCTACTCACCAACAAGATGC 526
Qy 1419 TGGAGTAATTTGTTGATAGGAGATCTGGACCTAAGGCTTTGTGGGGCTCATAGCCC 1478
Db 527 TGGAGTGACCTGCTCAGATGGATCCAATTTGGAAATGAGGCTGACGCGTGGAGGGAATAT 586
Qy 1479 CTGTTATGGGAGATGGAGGTGAATACCAAGAGAGTGGGGGACTGTGTCTATGACAG 1538
Db 587 GTGTTCTTGGAAAGATAGAGATCAAAATCCAAAGGAGCGTGGGGAACAGTGTGTGATGATA 646
Qy 1539 ATGGAGCACAAAGGATGAGCTGTGTGTGTPAAACAATTTGGGATGTGGAAGCCCTATGCA 1598
Db 647 CTCAACATAGATCATGCTGTCTGTTGTAGACAACCTGAATGTGGAAGTGTGTCAG 706
Qy 1599 TGTGTTGGTATGACCTATTTAAAGAGCATCAGGACCTATTTGGCTGGATGAGCTTTC 1658
Db 707 TTTCTCTGTTTCATCTAATTTTGGAGAGGCTCTGGACCAATCTGGTTGTGATGATCTTAT 766
Qy 1659 TTGATTGGAAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGGGGAAGACATAA 1718
Db 767 ATGCAACGGAATGAGTCAAGTCTCTGGAACCTGCAACATCAAGATGGGGAAGACATAA 826
Qy 1719 TTGTTACACAGAGAGGATGTGATTTGAACCTGCTCAGGTGATGCAACATGGGCGCTGAG 1778
Db 827 CTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAGGGAGCAGATCTGAGCCTCAG 886
Qy 1779 CTGTTGGGCGGACAGCAACCCCTGCTCGGGAAGACTGAGAGTGTACTTTTCAAGGACGGTG 1838
Db 887 ACTGTAGATAGGATCACTGAATGTTCAAGGAAGATTAGAAGTGGATTTCCAAGGAGAATG 946
Qy 1839 GGGCACAGTGTGATGACGCTGGAACAGTAAAGCTCCAGCTGTGGTGTGTAGCCAGCT 1898
Db 947 GGGACAATATGTGATGACGCTGGGACAGTTACGATGCTGTGGCATGCAAGCAACT 1006
Qy 1899 GGACTGCCCATCTTCTATTCATTGGCATGGGTCTGGGAACCGCTTACAGGATATGGAAA 1958
Db 1007 GGGATGTCCAACTCCGCTCAGACCATTTGGTTCGAGTTTAAAGCCAGTAAAGGATTTGGACA 1066
Qy 1959 AATTGGCTCGATGATGTTTCTGTGATGGAGATGAGTCAAGATCTCTGCTCATGCAAGAA 2018
Db 1067 CATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGTGTCTGCTGCAATGTAACA 1126
Qy 2019 CAGTGGGTGGGGAATATGACTGCAGTCAAGTGAAGATGTTGGAGTGTATCTGTCTGA 2078
Db 1127 CCATGAATGGGGAAGCATTTATGCAATCAAAATGAAGATGCTGGCGTGACATGTTCTGA 1186

Qy 2079 TGCATCGGATATGAGGCTGAGGCTTGTGGTGGAGCAGCAGGCTGTCTGGAAGAGTTGA 2138
Db 1187 TGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGCAGCCGCTGTCTGTGGACAGTTGA 1246
Qy 2139 GGTGAATGTCAGGCTGCGGCTGGGAATTTCTGTGTCTAATGGCTGGGAATGAACATTCG 2198
Db 1247 GCTGAGATTCAGAGACTGTTTAGGGAAGTGTGTGACAGAGGCTGGGACTGGAAGAAGC 1306
Qy 2199 TGAAGTGTGTTTTCAGGCAACTTGAATGCTGCTCAATCAGGCTCCAGAGAGCCCTCA 2258
Db 1307 TGTGTGTTTTGCGAGCAGCTGGGATGTGGATCTGCATCAAAACATCTTATCAAGTGA 1366
Qy 2259 TTTTCACAGAAAACATTTACACATCTTAATGTGCAATTTCTGGCTCAGCTGGAGGGGAAGC 2318
Db 1367 CTCAAAATCCAGGCAACAACACATGCTGTTTCTAAGTAGCTCTAAGCGGAATGAAC 1426
Qy 2319 CTCCTCTCTGGATTTGTATACGATGGGAGTGGAAACAGACTGGCTGTCTATTTAAATATGA 2378
Db 1427 TTTCTTTTGGGACTGCAAGAACTGGCAATGGGCTGGACTTACCTGTGATCACTATGAAGA 1486
Qy 2379 AGCAAGTTTGTATCTGCTCAGCCACAGCAGCCAGGCTGGTTGGAGCTGATATGCCCTG 2438
Db 1487 AGCCAAAATTTACCTGCTCAGCCACAGGGAACCCAGACTGGTTGGAGGGACATTTCCCTG 1546
Qy 2439 CTCTGGAGCTGTGAAAGTGAACATGCAGACACATGGCGCTCTGTCTGTGATCTGATTT 2498
Db 1547 TTTCTGGAGCTGTGAAAGTGAAGCATGTTGACACGTGGGCTCCATCTGTGATTCGGACTT 1606
Qy 2499 CTCCTCTCATGCTGCCAATGTGCTGTGCAGAGAAATTAATTTGTGGAGATGCCATATCT 2558
Db 1607 CTCTCTGGAAGCTGCCAGCGTTCTATGCAAGGAATTTACAGTGTGCACAGTTGTCTCTAT 1666
Qy 2559 TTTCTGTGGAGATCACTTTTGGAAAAGGAATGGTCTAATTTGGCGCGAAAAGCTTCCAGTG 2618
Db 1667 CTGGGGGAGACTCACTTTGGAGGGAATGGACAGATCTGGGCTGAAGATTTCCAGTG 1726
Qy 2619 TGAAGGAGTGAACCTCACTTTGCAATATGCCCCATTTTCAACATCCCGGAAGACACTTG 2678
Db 1727 TGAGGAGATGAGTCCCATCTTCTACTCTGCCAGTAGCACCCGCCCAAGAGAACTTG 1786
Qy 2679 TATCCACAGCAGAGAAATTTGGAGTTGCTGTGCCGATATACAGATGTCGGACTTGTGAA 2738
Db 1787 TAGCCACAGCAGGAGTGTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTGGTGA 1846
Qy 2739 TGGCAATCCCACTGTGACGGCAAGTGGAGATCAACGTGTTGGACACTGGGGCTCACT 2798
Db 1847 TGGCAAGACCCGCTGTGAGGCGAGAGTGSAGCTCAAAAGCTTTGGTCTGGGATCCCT 1906
Qy 2799 GTGTGACACCCACTGGGACCCAGAAAGTGCCTGTCTATCCACAGACTCAGCTGTGG 2858
Db 1907 CTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAATGTGG 1966
Qy 2859 GACTGCTCTCAACACAGAGGAGAAATATATTGAGAAAGAAAGTGTGCTGTGGGG 2918
Db 1967 AGTTCCCTTTCTACCCAGGAGGAGCAGGTTTTTGGAAAGGAATTTGTCAGATCTGGAG 2026
Qy 2919 ACACAGGTTTCACTTTAGGGAATGAGTCACTTCTGGAATACTGTCAAAATGACAGTTCT 2978
Db 2027 GCATATGTTTCACTGCACTGGGACTGAGCAGCAGATGGGAGATTTGCTGTAACTGCTCT 2086
Qy 2979 TGGAGCACTCCCTGTATCCATTCGAATACTCTCTCTGTGATCTGCACAGAGGCTGAC 3038
Db 2087 AGGTGCTTCAATTTATCTCTTTCAGAGCAAGTGGCCTCTGTAAATCTGCTGCAAGAAACAGTC 2146
Qy 3039 CCAGCCACTGTTTCCATGCTCCGCAATGTATCTGACCCATATTTGTGTGAGTTTCCAGA 3098
Db 2147 CCAACACACTGCTCTGTCGAATTTCACTGCTTTTGGGCCCAACAAGGCCCTTACCATTCCAGA 2206
Qy 3099 GGGCAGCTTGTGATCTGCTTAGAGSACAACCGCTCCGCTAGTGGATGGGAGGACAGCG 3158
Db 2207 AGAAAGTGTCTGGCCTGCATAGAGAGTGTGCTCACTTCCCTGGTAAATGGAGAGGTGG 2266
Qy 3159 CTGTCCCGGAGAGTAGAGATCTATACAGACGGCTTCTGGGGCAGCATCTGTGATGACGG 3218

[illegible]

Db	3344	AGGAGAGAAGCTT	TAGTCCACCAAAATTC	AAATACCGGAGATGAAT	TTCTTCGCTGAATGCAGA	3430
Qy	4281	GGACCCACATGGGACAAGAACCT	GATGATGAC	4311		
Db	3404	TGATCTGGACCTAATGAAT	TTCTCTCAGGAGGC	3434		
RESULT 7						
AAI60435/c						
ID	AAI60435	standard;	cDNA;	3834	BP.	
XX	AC	AAI60435;				
XX	XX	22-OCT-2001	(first entry)			
XX	DE	Human polynucleotide	SEQ ID NO 4424.			
XX	XX	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;				
KW	KW	peripheral nervous system; neuropathy; central nervous system; CNS;				
KW	KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;				
KW	KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;				
KW	KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;				
KW	KW	leukaemia; ss.				
OS	OS	Homo sapiens.				
XX	XX	WO200153312-A1.				
XX	PD	26-JUL-2001.				
XX	PF	26-DEC-2000; 2000WO-US34263.				
XX	XX	21-JAN-2000; 2000US-0488725.				
PR	PR	25-APR-2000; 2000US-0552317.				
PR	PR	09-JUL-2000; 2000US-0598042.				
PR	PR	13-AUG-2000; 2000US-0620312.				
PR	PR	03-AUG-2000; 2000US-0653450.				
PR	PR	14-SEP-2000; 2000US-0662191.				
PR	PR	19-OCT-2000; 2000US-0693036.				
PR	PR	29-NOV-2000; 2000US-0727344.				
XX	XX	(HYSE-) HYSEQ INC.				
PA	PA	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;				
PI	PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;				
PI	PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;				
XX	XX	WPI: 2001-442253/47.				
DR	DR	P-PSDB: AAM41279.				
XX	XX	Novel nucleic acids and polypeptides, useful for treating disorders				
PT	PT	such as central nervous system injuries -				
XX	XX	Claim 1; SEQ ID NO 4424; 10078pp; English.				
PS	PS	The invention relates to human nucleic acids (AAI57798-AAI61369) and				
XX	XX	the encoded polypeptides (AAM38642-AAM42213) with nootropic,				
CC	CC	immunosuppressant and cytostatic activity. The polynucleotides are used				
CC	CC	in gene therapy. A composition containing a polypeptide or polynucleotide				
CC	CC	of the invention may be used to treat diseases of the peripheral nervous				
CC	CC	system, such as peripheral nervous injuries, peripheral neuropathy and				
CC	CC	localised neuropathies and central nervous system diseases, such as				
CC	CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic				
CC	CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the				
CC	CC	utilisation of the activities such as: Immune system suppression,				
CC	CC	activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic				
CC	CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening				
CC	CC	assays for receptor activity, arthritis and inflammation, leukaemias and				
CC	CC	C.N.S disorders.				
CC	CC	Note: The sequence data for this patent did not form part of the print				
CC	CC	specification.				
XX	XX	Sequence 3834 BP; 979 A; 1082 C; 756 G; 1017 T; 0 other				

Query Match	31.3%	Score 1448.2	DB 22	Length 3834	
Best Local Similarity	65.4%	Pred. No. 0			
Matches 2179	Conservative	0	Mismatches 1128	Indels 24	Gaps 3
Qy 1002	GCTTGATGGTCTCCCTGCCGGTAATGAATCTTTCTTTGGGACTGCACACATCCGG	1061			
Db					
Db 3728	GCTACTTGAAGACTCTGGATCTGCTGACTTCAGAAGACATTTTGTCAACCTGAGTCCCTT	3669			
Qy 1062	AACCGTCAATTTTGACTGTCTTTCAATCAAAAGATGTGTCTGTGATCTGCTCAGATGAGC	1121			
Db					
Db 3668	CACCATTAAGTGTCTTACTTCTTCAGTGCCTGTTTGTCTACCAAGTCTCTCTTGAGGAAC	3609			
Qy 1122	AGATTGGAATCGGACTAGCAGATGGAAGTAAACAATTTGTTACGGAGAGTAGAGGTGAG	1181			
Db					
Db 3608	AGACAAGAGCTGAGCTAGTGGATGGTGAANAACATGTATACGGGAGAGTGGAGGTGAA	3549			
Qy 1182	AAATTCAATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGT	1241			
Db					
Db 3548	AGTCCAGAGGAGTGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTCTCTGT	3489			
Qy 1242	GGTTTGAACAGCTAGGATGTCGGTTCAGCGCTTTTGGCAGTCTGTCGTCTAAACCTAG	1301			
Db					
Db 3488	GATTTGAACAGCTGGGATGTCCAACCTGCTATCAAAAGCCCTGGATGGGCTAAATCCAG	3429			
Qy 1302	TAATGAAGCTAGACACATTTGGATAAACAGACATATCTTGCACTGGGAATGAGTCAGCTCT	1361			
Db					
Db 3428	TGCAGGTTCTGGAGCATTTTGGATGGATCATGTTTCTGTGCTGGGAATGAGTCAGCTCT	3369			
Qy 1362	CTGGGACTGCACATATGATGGAAGCAAGCAAGC---GAACATGCTTCCGAAGATCAGATGC	1418			
Db					
Db 3368	TTGGGATGGAACATGATGGATGGGAAGCAATAGTAACCTGTACTACCAACAAGATGC	3309			
Qy 1419	TGGAGTAATTTGTTCTGATAAGGCAGATCTGCACCTAAGCGTGTGCGGGGCTCATAGCCC	1478			
Db					
Db 3308	TGGAGTGACCTGCTCAGATGATCCAATTTGGAATGAGGCTGACCGTGGAGGAATAT	3249			
Qy 1479	CTGTTATGGGAGATGGAGGTGAAATCCAAGGAGAGTGGGGGACTGTGTCTCATGACAG	1538			
Db					
Db 3248	GTGTTCTGGAAGAATAGAGATCAAAATTCGAAGGAGCGTGGGGAACAGTGTGTGATGATA	3189			
Qy 1539	ATGGAGCACAAGGAATGCAGCTGTTGTGTGTAAACAATTTGGGATGTGGAAGCCCTATGCA	1598			
Db					
Db 3188	CTTCAACATAGATCATGCATCTGTCTATTTGTAGACAACTTGAATGTGGAAGTGTGTGAC	3129			
Qy 1599	TGTTGTTGGTATGACCTATTTTAAAGAACATCAGGACCTATTTGGCTGGATGACGTTTC	1658			
Db					
Db 3128	TTTCTCTGGTTCACTAATTTTGGAGAAGGCTCTGGACCAATCTGGTTGATGATCTTAT	3069			
Qy 1659	TTGCATTTGGAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGGGGAAGCATAA	1718			
Db					
Db 3068	ATGCAACGGAATGAGTCAAGCTCTCTGGAACCTGCAAAACATCAAGGATGGGGAAGCATAA	3009			
Qy 1719	TTGTCTACACAGAGGATGTGATTGTAACTGCTCAGGTGATCAACATCGGGGCTGAG	1778			
Db					
Db 3008	CTGTGATCATCTGAGGATGCTGGAGTGATTTGCTCAAGGGAGCAGATCTGACGCTGAG	2949			
Qy 1779	GCTGGTGGGCGGACAAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTG	1838			
Db					
Db 2948	ACTGGTAGATGGAGTCACTGAATGTTCAAGGAAGATTAGAAGTGAAGTATCCAAGGAATG	2889			
Qy 1839	GGGCACAGTGTGATGACGCTGGAACAGTAAAGCTCAGCTGTGGTGTCTAGCCAGCT	1898			
Db					
Db 2888	GGGACAATATGTGATGACGCTGGACAGATTACGATGCTGCTGTGGCATGCAAGCAACT	2829			
Qy 1899	GGACTGCCCATCTTCTATCATTTGGCATGGGTCTGGGAACGCTTCTACAGGATATGAGAA	1958			
Db					
Db 2828	GGGATGTCCAACCTGCCGTACAGCAATTTGGTCCAGTTAAACGCCAGTAAGGATTTGGACA	2769			
Qy 1959	AATTTGGCTCGATGATGTTTCTGTGATGGAGATGAGTCAGATCTCTGGTCATGCAAGAA	2018			
Db					
Db 2768	CATCTGGCTTGACAGCGTTTCTTGCCAGGGAGATGAACCTGCTGTCTGGCAATGTAAACA	2709			

Db 1628 AGAAGTCTGTGGCTGCATAGAGAGTGTCAACTTGCCTGGTAAATGGAGAGGTGC 1569
QY 3159 CTGTGGCGGAGAGTAGAGATCTATCAGGACGGCTTCTGGGGCACCATCTGTGATGACGG 3218
Db 1568 CTGTGTGGGAGAGTAGAGATCTATCATGAGGCTCTCTGGGGCACCATCTGTGATGACAG 1509
QY 3219 CTGGGACCTCAGCGATGCCACACCTGTGTGTCAAAAGCTGGGTGTGGAGTGGCCTTCAA 3278
Db 1508 CTGGGACCTCAGTGATGCCACCTGTGTGTCACACAGCTGGGTGTGGAGAGCCATTAA 1449
QY 3279 TGGCACGGTCTGTCTCAGCTTTTGGGAGGGGTGAGGGCCCATCTGGCTGGATGACCTGAA 3338
Db 1448 TGGCACTGGTCTGTCTCATTTGGGAAGAACAGAGGCCCATCTGGCTGGATGAGATGAA 1389
QY 3339 CTGCACAGGAACGAGTCCACCTGTGTGGCAGTGCCTTCCCGCGGTGGGGCAGCAGCA 3398
Db 1388 ATGCAATGGAAAGATCCCGCATTTGGCAGTGCCATTCACAGGCTGGGGGAGCAAAA 1329
QY 3399 CTGCAGGCACAAAGGAGGACGAGGGTCTATCTGCTCAGAAATTCACAGCCCTTGAGGCTCTA 3458
Db 1328 TTGCAGSCACAAGGAGGATCGGGAGTTATCTGCTCAGAAATTCATGCTCTGAGACTGAC 1269
QY 3459 CAGTGAACCTGAACACAGAGAGCTGTGCTGGAGATTTGGAAGTCTTCTATAACGGGACCTG 3518
Db 1268 CAGTGAAGCCAGCAGAGAGAGGCGCTGTGCAGGCGCTGTGGAAGTCTTTTACAAATGGAGCTTG 1209
QY 3519 GGGCAGGCTCGGAGGAGGAACATCACACAGCATAGCAGGATTTGTGCAGGAGGCT 3578
Db 1208 GGGCAGCTTTGGCAGAGTAGCATGTCTGAACACACTGTGGGTGTGTGTCAGGAGCT 1149
QY 3579 GGGCTGTGGGAGAAATGAGTGTGTAGCTCTGCCCTCTTTATCTAAGACAGGCTCTGGTTT 3638
Db 1148 GGGCTGTGCAGACAAAGGAAATCAACCTGTCTTTAGACAAGGCCATGTCCATTCC 1089
QY 3639 CATGTGGTGGATGACATTCAGTGTCTCTAAACGCATATCTCCATATGGCAGTGCCTGTC 3698
Db 1088 CATGTGGTGGACAATTTTCAGTGTCTCAAAAGGACCTGTGACACGCTGTGGCAGTGCCTCATC 1029
QY 3699 TGCCCCATGGGAGGAGAAATCTCCAGCCAGCAGAGACCTTGGATCATCATGTGAAGA 3758
Db 1028 ATCTCCATGGGAGAGAGACTGGCCAGCCCTCGGAGGAGACCTGTGATCATCATGTGACAA 969
QY 3759 TAGAATAAGAGTGGTGGAGAGACACCGAGTGTCTGGAGAGTGGAGATCTGGCAGCG 3818
Db 968 CAAGATAAGACTTCAGGAAGGCCACCTCTCTGTCTGGAGCTGTGGAGATCTGGCAGTG 909
QY 3819 AGGCTCTGGGACACAGTGTGTGATGACTCTGGGACCTGGCCAGCGGAGTGTGTG 3878
Db 908 AGGTCTCTGGGGACAGTGTGTGATGACTCTGGGACTTGGACGATGCTCAGGTGTGTG 849
QY 3879 TCAGCAGCTGGGCTGTGGCTGTGCTGTGGCTGGCCCTGAGGAGCGCTTCGTTTGGCCAGGG 3938
Db 848 TCAACAACCTGGCTGTGGTCCAGCTTTGAAGACATTCAAAGAACGAGATTTGTCTCAGGG 789
QY 3939 AACTGGAAACATCTGGTGGATGACATGCGGTGCAAGAGAAATGAGTCATTTCTATGGGA 3998
Db 788 GACTGGACCATATGGCTCAATGAAGTGAAGTGCAAAGGAATGAGTCTTCTTGTGGGA 729
QY 3999 CTGTACGCCCAACCCCTGGGACAGAGTGTGTGGACACAAGGAAGATCTGCGCGTAG 4058
Db 728 TTGTCTGCCAGACGCTGGGGCCATAGTGTGTGGGACAGAGAACCCCTGCAGTGAA 669
QY 4059 GTGCTCTGGACAGTGCCTGAA-----ATCACTGAATGCCCTCCTCAGG 4100
Db 668 TTGCACAGATATTTCACTGCAGAAAACCCACAAAAGCCACAAACAGGTCGCTCATCCCG 609
QY 4101 TCATTTAGCACTTATTTTATCCAGTATCTTTGGCTCCTCTCTCTGTTGTTGTTTATCT 4160
Db 608 TCAGTATCCTCTTATTCAGTTCGGGATCTCTGGGTTGTTCTGTGTGGCCATTTTCGTGCG 549
QY 4161 ATTTCTCAGCTGGCGGAGTTTCAAGAAACAAACATCTCCCTCAGAGTTTCAACAG 4220

Db 548 ATTATTTCTTCTTGACTAAAGGAGAGACAGACACA---CGGCTTGCAGTTTCTCTCAAG 492
QY 4221 AAGGAGGGGTCTCTCTCGAGGAGAAATTTATTCATGAGATGGAGACCTTGCTCAAGAGAGA 4280
Db 491 AGGAGAGAACTTAGTCCACCAATTCATACCGGAGATGAATTTCTTGCCTGAATGCAGA 432
QY 4281 GGACCCACATGGGACAAAGAACCTTCAGATGAC 4311
Db 431 TGATCTGGACCTAATGAATTCCTCAGGAGGC 401

RESULT 8
AAI60436/c
ID AAI60436 standard; cDNA; 3834 BP.
XX AAI60436;
AC AAI60436;
XX 22-OCT-2001 (first entry)
DT Human polynucleotide SEQ ID NO 4425.
DE
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX Homo sapiens.
OS
XX WO200153312-A1.
PN
XX 26-JUL-2001.
PD
XX 26-DEC-2000; 2000WO-US34263.
PF
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-052317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR P-PSDB; AAM41280.
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 4425; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-brager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed
specification.

XX
SQ Sequence 3834 BP; 979 A; 1082 C; 756 G; 1017 T; 0 other:
Query Match 31.3%; Score 1448.2; DB 22; Length 3834;
Best Local Similarity 65.4%; Pred. No. 0;
Matches 2179; Conservative 0; Mismatches 1128; Indels 24; Gaps 3;

QY 1002 GCTTCATGCTCTCTCTCGGTAATGAATCTTTCTTTGGGACTCAGACATTCGGG 1061
DB 3728 GCTACTTGAAGACTCTGGATCTGCTGACTTCAGAGACATTTTGTCAACCTGAGTCCCTT 3669
QY 1062 AACCTCAATTTGACTGTCTTCATCAAAACGATGTCTGTGATCTGCTCAGATGGAGC 1121
DB 3668 CACCATTACTGTGCTTACTTCTCAGTGCCTGTTTGTACCACGTTCTCTTGGAGAAC 3609
QY 1122 AGATTTGACACTGACATAGCAGATGGAAGTAACAAATTTTCCAGGAGAGTAGAGTGAG 1181
DB 3608 AGACAAGGAGCTGAGGCTAGTGATGGTGAAGTCAAAACAAAGTGTAGCGGAGAGTGAAGTGAA 3549
QY 1182 AATTCATGAACAGTGTGGACAATATGTACCAGACACTGGAGAAATGAACAGCCCTTGT 1241
DB 3548 AGTCCAGGAGGAGTGGGAACGGTGTGAATAATGGCTGGAGCATGGGAAGCGGTCTCTGT 3489
QY 1242 GCTTTGTAAGCAGTAGGATGCCCTTTCAGCGCTCTTTGGCAGTCTGCTGCTFAAACCTAG 1301
DB 3488 GATTTGTAACAGCTGGGATGTCCTCAACTGCTATCAAGCCCTGGATGGCTTAATCCAG 3429
QY 1302 TAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACCTGGGAATGAGTCACTCT 1361
DB 3428 TGCAGGTTCTGACGCATTTGGATGATCATGTTCTTGTCTGGGAATGAGTCACTCT 3369
QY 1362 CTGGGACTGCACATATGATGGAAGCAAGC ---GAACATGCTTCCGNAATCAGATGC 1418
DB 3368 TTGGGATTGCAACATGATGATGGGAAGCATAGTAACCTACTACCAACAAAGATGC 3309
QY 1419 TGGAGTAATTTGCTGTAAGGCAGATCTGGACCTAAGGCTGTGCGGGCTCATAGGCC 1478
DB 3308 TGGAGTACCTCTCAGATGGATCCAAATTTGGAAATGAGCTGACGCGTGGAGGAATAT 3249
QY 1479 CTGTTATGGAGATTTGGAGTGAATACCAAGGAGAGTGGGGATGTGTGTATGACAG 1538
DB 3248 GTGTTCTGGAAGATAGATCAAAATTCGAAGCGGTGGGAACAGTGTGTGATGATAA 3189
QY 1539 ATGGAGCAAGGATGAGCTGTGTGTGAACAATTTGGATGTGGAACCCATATGCA 1598
DB 3188 CTCAACATAGATCATGATCTGTCTATTTGTAGACAACCTTGAATGTGGAAGTGTGTCAG 3129
QY 1599 TGTGTTTGGTATGACCTATTTTAAAGACATCAGGACCTATTTGGCTGGATGAGCTTTC 1658
DB 3128 TTTCTCTGTTCTATCTAATTTTGGAGAGGCTCTGGACCAATCTGTTTGTATGATCTTAT 3069
QY 1659 TTGCATTTGGAATAGTCAAAATATCTGGGACTGTGAACACAGTGGATGGGGAAGCATAA 1718
DB 3068 ATGCAACGGAATAGTCAAGCTCTCTGGAACCTGCAACATCAAGATGGGGAAGCATAA 3009
QY 1719 TTGTGTACACAGAGGATGTGATTTAACTGCTCAGGTGATGCAACATGGGGCTGAG 1778
DB 3008 CTGTGATCATGTGAGGATGTGGAGTGTATTTGCTCAAGGGAGCAGATCTGAGCGCTGAG 2949
QY 1779 GCTGTTGGGCGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTCAAGGACGGTG 1838
DB 2948 ACTGTTAGTATGAGTCACTGAAATGTTTCAAGAAAGATTAAGAAATGAGATTTCCAAGGAGATG 2889
QY 1839 GGGCAGCTGTGTATGAGCGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCT 1898
DB 2888 GGGGCAATATGTGATGACGGCTGGACAGTTACGATGCTGTGTGGCATGCAAGCAACT 2829
QY 1899 GGACTGCCCATCTTCTATCATTTGGCATGGGCTGGGAACGCTTCTACAGGATATGGAA 1958
DB 2828 GGGATGTCCAACTGCCGTCACAGCCATTTGGTCGAGTTAAGCCAGTAAGGGATTTGGACA 2769

QY 1959 AATTTGGCTCGATGATGTTTCTCTGTGATGGAGATGAGTCAGATCTCTGTCATGCAGAA 2018
DB 2768 CATCTGGCTTGACACCGTTTCTTGGCAGGACATGAACCTGCTGTCTGCAATGTAAACA 2709
QY 2019 CAGTGGGTGGGAAATTAATGACTGCAGTCACAGTGAAGATGTTGGAGTATGTTCTGGA 2078
DB 2708 CCATGAATGGGAAACGATTTATTTGCAATCACAATGAAGATGCTGGCGTGACATGTTCTGA 2649
QY 2079 TGCATCGGATATGGAGCTGAGGCTTGTGGTGAAGCAGCAGCTGTCTGGAAGAAATGGA 2138
DB 2648 TGGATCAGATCTGGAGCTTAAGACTTAGAGGTGGAGCAGCCGCTGTCTGCGACAGTGA 2589
QY 2139 GGTGAATGCCAGGTCGCCGGGAATTTCTGTGCTCTAATGGCTGGGAATGAACATTCG 2198
DB 2588 GGTGAGATTCAGACACTGTTAGGCAAGCTGTGTGACAGAGGCTGGGACTCGAAGAAGC 2529
QY 2199 TGAAGTTGTTGCCAGGCAACTTGAATGTGGTCTCAATCAGGCTCTCCAGAGACCTCA 2258
DB 2528 TGATGTGTTTGCAGGCAAGCTGGGATGTGATCTGCACCTCAAAACATCTTATCAAGTGA 2469
QY 2259 TTTTCACAGAAACATTTACACATCTTAATGTGCAATTTCTGCTGCACCTGGAAGCAAGC 2318
DB 2468 CTCCAAAATCCAGGCAACAAACACATGGCTGTTTCTAAGTAGCTGTACGCGAAATGAAC 2409
QY 2319 CTCTCTCTGGGATTTATACGATGGGAGTGGAAACAGACAGCTGCTGTCTATTTAAATATGA 2378
DB 2408 TTTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTATGAAG 2349
QY 2379 AGCAAGTTTGTCTCTCAGCCCAAGGAGCCAGGCTGTGGAGCTGTATGCCCTG 2438
DB 2348 AGCCAAAATTTACCTCTCAGCCCAAGGAGAACCCAGACTGTTGGAGGGGACATTTCCCTG 2289
QY 2439 CTCTGAGCTGTGAGTGAACATGCACACATGGCGCTCTGTCTGTGATTTGATTT 2498
DB 2288 TTTCTGAGCTGTGAGTGAAGCTGAGCAGTGGTGGGCTCCATCTGTGATTTGGGACTT 2229
QY 2499 CTCTCTTCTGCTGCAATGTCTGTGCAGAGAAATTAATTTGGAGATGCCATATCTCT 2558
DB 2228 CTCTCTGGAAGCTGCCAGGCTTCTATGCAAGGAATTTACAGCTGTGCACAGTTGTCTCTAT 2169
QY 2559 TTTCTGAGGATCACTTTTGGAAAGGNAATGCTTAACCTTGGCGCGGAAAGATTTCCAGTG 2618
DB 2168 CTTGGGGGAGCTCACTTTTGGAGAGGAAATGGACAGATCTGGGCTGGAAGAAATTTCCAGTG 2109
QY 2619 TGAAGGGAGTGAACCTCACTTTGCAATTTATGCCCATTTGTTCAACATCCGGAAGACACTTG 2678
DB 2108 TGAGGACATGAGTCCCATCTTTCACTCTGCCAGTAGGACCCGCCCAAGAGAACTTG 2049
QY 2679 TATCCACAGCAGAGAGTTGGAGTTGTCTGTTCCGATATACAGATGTCCGACTTTGTGAA 2738
DB 2048 TAGCCACAGCAGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTGGTGA 1989
QY 2739 TGGAAATCCAGTGTGAGCGCAAGTGGAGATCAACGTGCTTGGACACTGGGCTCACT 2798
DB 1988 TGGCAAGACCCGCTGTGAGGCGAGAGTGGAGCTCAAAAGCTTTGGTCTGCTGGGATCCCT 1929
QY 2799 GTGTGACACCCACTGGGCGGAGAGTCCGCTGCTTCTATCGACAGCTCAGCTGTGG 2858
DB 1928 CTGTAACCTCACTGGGACATAGAAGTGGCCATCTTTTGGCAGAGCTTTAAATGTGG 1869
QY 2859 GACTGCTCTCTCAACACAGAGGAAATATATTGGAGAAAGAGTGTTCGTGTGGGG 2918
DB 1868 AGTTGCCCTTCTACCCAGGAGGACGAGTTTGGAAAGGAATTTGTCAGATCTGGAG 1809
QY 2919 ACAGAGTTTCAATGCTTAGGGAATGAGTCACTTCTGGAAATGTCATAATGACAGTTCT 2978
DB 1808 GCATATGTTTCACTGCACTGGGACTGAGCAGCAGATGGGAGATTTGCTGTCACTGCTCT 1749
QY 2979 TGGACACCTCCCTGTATCCATGGAATACTCTCTCTGTGATGTCACAGAGGCTGAC 3038
DB 1748 AGGTGCTTCATTTATGCTTTCAGAGCAAGTGGCCCTGTAAATCTGCTAGGAAACAGTCT 1689
QY 3039 CCAGGCCACTGTTTCCATGCTCGCAAAATGATATCTGACCCATATTTTGTCTGCTGAGTCCAGA 3098

Db 107 GCTACTGAAGACTCTGGATCTGCTGACTTCAAGAAGACATTTTGTCAACCTGAGTCCCTT 166
QY 1062 AACCGTCAATTTTGAAGTCTTTCATCAAAACGATGTGCTGTGATCTGCTCAGATGAGC 1121
Db 167 CACCATTAATCTGCTTACTTCTCAGTCCCTGTTTGTCAACAGTCTCTTGGAGAAC 226
QY 1122 AGATTTGGAATCGGACTAGCAGTGAAGTAAACAATTTGT-TCAGGAGAGTAGAGGTGA 1180
Db 227 AGACAAGGAGCTGAGGCTAGTGGATGGTGAAGAACAGTGTAGCGGGAGAGTGAAGTGA 286
QY 1181 GAATTCATGAACAGTGGTGGACAAATATGTACCAGAACTGGAAGATGAACAGCCCTTG 1240
Db 287 AAGTCCAGGAGGAGTGGGACCGTGTGTAATATGGCTGGAGCATGGAAGGCTCTCTG 346
QY 1241 TGGTTTGAAGCAGTAGGATGTCGGTTCAGCGCTTTTGGCAGTCGTCGCTAAACCTA 1300
Db 347 TGATTTGAACAGTGGATGTCCTAACTGCTATCAAGCCCTTGGATGGCTTAATTTCCA 406
QY 1301 GTAATGAAGCTAGAGACATTTGGATTAACAGCATATCTTGCACCTGGGAATGAGTCAAGCTC 1360
Db 407 GTCAGGTTCTGGACGATTTGGATGGATGATGATGTTCTTCTGCTGGGAATGAGTCAAGCTC 466
QY 1361 TCTGGGACTGCATATGATGGAAGAACAAAGC---GAACATGTTCCGGAAGATCAGATG 1417
Db 467 TTTGGGATTGCAACATGATGGATGGGAAAGCATAGTAACCTGTACTCAACCAACAAGATG 526
QY 1418 CTGGAGTAATTTGTCTGATAAGCAGATCTGGACCTAAGCCTTGTGCGGSCCTCATAGCC 1477
Db 527 CTGGATGACCTGCTCAGATGGATCCAAATTTGGAATGAGGCTGACCGCTGGAGGGAATA 586
QY 1478 CCTGTTATGGAGATTGGAGGTGAATACCAAGGAGAGTGGGAGCTGTGTGTCATGACA 1537
Db 587 TGTGTTCTGGAAGAATAGACATCAAAATTCAGGACGCTGGGGAACAGTGTGTGATGATA 646
QY 1538 GATGGACCAAGGAATGCAAGCTGTGTGTGTAACAAATTTGGATGTGGAAAGCCTATGC 1597
Db 647 ACTTCAACATAGATCATGATCTGTCTATTGTTAGACAACTTGAATGGAAGTGTCTGCA 706
QY 1598 ATGTGTTGCTATGACCTATTTTAAAGAACATCAGGACCTATTGCTGCTGGATGACGCTT 1657
Db 707 GTTCTCTGTTCTAATTTTGGAGAGGCTCTGACCACTTGTGTTTGTGATGATCTTA 766
QY 1658 CTGCAATTTGGAATAGCTCAAAATATCTGGACATGTGAACAGTGGATGGGAAAGCATA 1717
Db 767 TATGCAACGGAATAGCTCAGCTCTCTGGAATGCAAAACATCAAGGATGGGAAAGCATA 826
QY 1718 ATGTGTACACAGAGGATGTGATGTAACTGCTCAGTGATGCAACATGGGCGCTGA 1777
Db 827 ACTGTGATCATGTGAGGATGCTGGAGTGTATTTGCTCAAAAGGAGCAGATCTGAGCGCTGA 886
QY 1778 GCGTGTGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGT 1837
Db 887 GACTGGTAGATGAGTCACTGAATGTTTCAAGAGATTTAGAAGTGAATTTCAAGGAGAA 946
QY 1838 GGGGACAGTGTGTGATGACGGCTGGAAACAGTAAAGCTGACGCTGGTGTGTAGCCAGC 1897
Db 947 GGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGTGTCGTCATGCAAGCAAC 1006
QY 1898 TGGACTCCCATCTTCTATCATTTGGCATGGTCTGGGAACCGTCTTACAGGATATGGA 1957
Db 1007 TGGGATGTCCAACTGCGCTCACAGCCATTTGGTGGAGTTAAGCCAGTAAGGGATTTGGAC 1066
QY 1958 AAATTTGGCTCGATGATGTTTCTGTCATGGAGATGAGTCAAGTCTGCTGATGAGCA 2017
Db 1067 ACATCTGGCTTGACAGGCTTTTTCGCCAGGACATGAACCTGCTGCTGSCAATGTAAC 1126
QY 2018 ACAGTGGGTGGGAAATAATGACTGCAAGTCAAGTGAAGATGTTGGAGTCACTGTGTTCTG 2077
Db 1127 ACCATGAATGGGAAAGCAATTTTGAATCACAATGAAGATGCTGCGCTGACATGTTCTG 1186
QY 2078 ATGCATCGGATATGAGGCTGAGGCTTGTGGTGGAAACAGCAGGATGTGCTGGAAAGTTG 2137
Db 1187 ATGGATCAGATCTGGAGCTAAGACTTAGAGTGGAGGACGCGCTGTGCTGGGACAGTTG 1246

QY 2138 AGTGAATGTCCAGGGTCCGCTGGGAATCTGTGTCTAATGTGCTGGGAATGAACATTTG 2197
Db 1247 AGTGGAGATTCAGAGACTGTTAGGGAAGTGTGTGACAGAGCTGGGAGCTGAAAG 1306
QY 2198 CTGAAGTGTGTTGCAGGCAAACTTGAATGTGGGTCTGCAATCAGGGTCTCAGAGAGCCT 2257
Db 1307 CTGATGTGGTTCGAGCGAGCTGGATGTGATCTGCACTCAAAACATCTTATCAAGTGT 1366
QY 2258 ATTTCAAGAAACACATTAACATCTTAATGTCAATTTGCTGCTGCACTGGAGGGAAG 2317
Db 1367 ACTCCAAAATCCAGGCAACAAACACATGCTGTTTCTTAAGTAGCTGTAAACGGAATGAA 1426
QY 2318 CCTCTCTCTGGGATTTGATACGATGGAGTGGAAACAGACTGCGGTGTCATTTAAATATGG 2377
Db 1427 CTCTCTCTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTATCAAG 1486
QY 2378 AAGCAAGTTTGAATCTGCTCAG--CCCACAGGACGCCAGGCTGGTGGAGCTCATATGCC 2435
Db 1487 AAGCAAAATTTACCTGCTCAGGCCACAGGGAAACCCAGACTGCTTGGAGGGGACATTC 1546
QY 2436 CTGCTCTGGAGCTGTTGAAGTGAACATGCAGACATGCGGCTCTCTCTGTGATTTCTGA 2495
Db 1547 CTGTTCTGGAGCTGTTGAAGTGAAGCATGTGTGACACGTGGGCTCCATCTCTGTGATTCAGA 1606
QY 2496 TTTCTCTCTCTATGCTGCCAATGTGCTGTGCAGAGAATTAATTTGTGGAGATGCCATATC 2555
Db 1607 CTCTCTCTGGAAGCTGCCAGCGTCTTATCGAGGGAATACAGTGTGGCAGCTGTCTC 1666
QY 2556 TCTTTCTGTGGGAGATCACTTTGGAAAGGAATGGTCTAACTTTGGGCCGAAAGATTCCA 2615
Db 1667 TATCTCTGGGAGCTCACTTTGGAGAGGAAATGGACAGATCTGGCTGAAGAAATTTCCA 1726
QY 2616 GTGTGAGGAGTCAAACTCACCTTGCAATATGCCCATTTTCAACATCCGGAAGACAC 2675
Db 1727 GTGTGAGGACATGAGTCCCCTCTTCTCACCTGCCCCAGTACACCCCGCCAGAGGAAC 1786
QY 2676 TTGTATCCACAGCAGAGAAGTTGGAGTTGCTGTTCCCGATATACAGATGCTCCAGCTTGT 2735
Db 1787 TTGTAGCCACAGCAGGAGTGTGAGTAGTCTGCTCAAGATACACAGAAATTCCTTGTGT 1846
QY 2736 GAATGGCAATCCAGTGTGACGGCAAGTGGAGATCAAGTGTGCTGGAACACTGGGCGCTC 2795
Db 1847 GAATGGCAAGACCCCGTGTGAGGCGAGTGGAGCTCAAAACGCTTGGTGCCTGGGATC 1906
QY 2796 ACTGTGTACACCCACTGGGACCCAGAGATGCCGTGTTCTATGACAGACAGCTCACGCTG 2855
Db 1907 CCTGTGAATCTCACTGGGACATAGAGATGCCCATGTTCTTTGGCAGCAGCTTAATG 1966
QY 2856 TGGGACTGCTCTCTCAACCCAGGAGGAAATATATTGGAGAAAGAAAGTGTCTGTGTG 2915
Db 1967 TGGAGTTGCCCTTTCTACCCAGGAGGAGCAGCTTTTGGAAAGAAATGGTCAAGATCTG 2026
QY 2916 GGGACAGAGTTTCAATGCTTAGGAAATGAGTCACTTCTGATTAACCTGTCAATGACAGT 2975
Db 2027 GAGSCATATGTTCACTGCACTGGGACTGAGCAGACATGGGAGATTTCTCTGTAAGTGC 2086
QY 2976 TCTTGGAGCAGCTCCCTGTATCCATGCAATACTGCTCTGTGATCTGCACAGGAAGCCT 3035
Db 2087 TCTAGTGTCTCATATGTTCTTCCAGAGCAAGTGGCCCTCTGTAATCTGCTCAGGAACCA 2146
QY 3036 GACCCAGCCACTGTTCCATGCTCGCAAAATGATCTGACCCATATTGTTGTCTGAGTTC 3095
Db 2147 GTCCAAACACTGTCCCTCGTGAATTCATGCTCTTGGGCCCAACAAGGCTACCATTC 2206
QY 3096 AGAGGAGTGTCTTGTGATCTGCTTAGAGGACAAACGGCTCGGCCCTAGTGGATGGGACAG 3155
Db 2207 AGAAGAAGTCTGTGCGCTGCAATAGAGAGTGTGTAACATTTGCGCTGGTAAATGGAGAGG 2266
QY 3156 CCGCTGTGCGGAGAGTAGAGATCTATCAGGAGCGCTTCTGGGGACCACTCTCTGATGA 3215
Db 2267 TCGCTGTGTGGGAGATAGAGATCTATCATGAGGCGCTCTCTGGGGGCCACCATCTCTGTGATGA 2326

Query Match	28.9%	Score 1339.2;	DB 22;	Length 3811;
Best Local Similarity	63.5%	Pred. No. 0;		

QY 4277 GAGAGGACCCACATGGGACAAGAACCCTCAGATGAC 4311

Matches 2179; Conservative 0; Mismatches 1128; Indels 123; Gaps 4;

QY	1002	GCTTGATGGTGTCTCCTCGTCGAGTAATGAATCTTTCTTTGGACATGCAGACATCCCGG	1061
Db	107	GCTACTTGAAAGACTCTGGATCTGCTGACTTCAGAGAACAATTTTGTCAACTGAGTCGCCCT	166
QY	1062	AACCGCTCAATTGTTGACTGTCTTCATCAAACAGATGTCTGTGATCTGCTCAGATGGAGC	1121
Db	167	CACCATTACTGTGCTTACTTCTCAGTGCCCTGTTTTGTCCACAGTTCTCTTTGGAGGAAC	226
QY	1122	AGATTGGNACTGCGACTAGCAGATGGAAGTAACAATTTTTCAGGGAGACTAGAGGTGAG	1181
Db	227	AGACAAGGAGCTGAGGCTAGTGGATGTGTAACACAAGTGTAGCGGAGAGTGGAAAGTAA	286
QY	1182	AATTCATGAACAGTGGTGGACAAATATGTACCAGCAACTGGNAGAAATCAACAAGCCCTTGT	1241
Db	287	AGTCAGGAGGAGTGGGAAACGGGTGTGAATAATGGCTGGAGCATGGAACGGTCTCTGT	346
QY	1242	GGTTTGTAAAGCACTAGGATGCCCTTCAGCGCTCTTTGGCAGTCTGCTGCTTAAACCTAG	1301
Db	347	GATTTGTAAACCACTGGGATGTCCAAGTGCTATCAAGGCCCTGGATGGGCTAATTCAC	406
QY	1302	TAAAGAAGCTAGAGACATTTGGATAAACAGATATCTTTGCACCTGGGAATGAGTCAGTCT	1361
Db	407	TGCAGGTTCTGCAGCAATTTGGATGGATCATGTTTCTTGCTGGGAATGAGTCAGCTCT	466
QY	1362	CTGGGACTGCACATATGATGGAAAAGCAAGC --- GAACATGCTTCCGAAGATCAGATGC	1418
Db	467	TTGGGATTGCCAACATGATGATGGGAAAGCATAGTAACCTGTACTCACCAACAAGATGC	526
QY	1419	TGGAGTAATTTCTTCTGATAGGCGAGATCTGGACCTTAGGCTTGCTGGGCTCATAGCCC	1478
Db	527	TGGAGTGACCTGCTCAGATGGATCCATTTTGGAAATGAGGCTGACGCGTGGAGGAATAT	586
QY	1479	CTGTTATGGAGATTGGAGGTGAATACCAAGGAGAGTGGGGGACTGTGTTCATGACAG	1538
Db	587	GTGTTCTGGAAGATAGAGATCAAAATCCAAAGACGGTGGGCAACAGTGTGATGATAA	646
QY	1539	ATGGAGCAAGAAGATGACGTGTGTGTGTAAACAATTTGGATTTGGAAAGCCTATGCA	1598
Db	647	CTTCACATAGATCATGATCTGTCATTGTGTAGACAATTTGAATGTGGAAGTGTGTCAG	706
QY	1599	TGTGTTTGGTATGACCTATTTTAAAGAACATCAGGACCTATTTGGCTGGATGACGTTTC	1658
Db	707	TTTCTCTGGTTCACTAATTTTGGAGAGCTCTGGACCAATCTGTTTGTATGATCTAT	766
QY	1659	TTGCATTGGAATAGTCAAAATATCTGGGACTGTGAACACAGTCGGATGGGGAAGCATAA	1718
Db	767	ATGCACGGAAATGATGACGTCTCTCGAACTGCAAAATCAAGGATGGGGAAGCATAA	826
QY	1719	TTGTGTACACAGAGGATGTGATTTAACTGCTCAGGTGATCCAACATGGGCGCTGAG	1778
Db	827	CTGTGATCATGCTGAGGATGCTGGAGTGATTTGTCTAAAGGGAGCAGATCTGAGCCCTGAG	886
QY	1779	GCTGTGGCGCGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTTCAAGGACGGTG	1838
Db	887	ACTGGTAGATGGATCACTGAATGTTTCAGGAAGATTAGAAGTGAGATTTCAAGGAGATG	946
QY	1839	GGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGTGTGTAGCCAGCT	1898
Db	947	GGGCAAAATATGTGATGACGGCTGGGACAGTTACGATGCTGTGGCATGCAAGCAACT	1006
QY	1899	GGACTGCCCATCTTCTATCATTTGGATGGGTCTGGGAACCCCTTCTACAGGATATGAAA	1958
Db	1007	GGATGTCCAATGCCGTACAGCCATTGGTCGAGTTTAAACCCAGTAAAGGATTTGGACA	1066
QY	1959	AATTTTGGCTCGATGATTTTCCCTGTGTAGTGAGATGAGTCAGATCTCTGGTTCATCGAGAA	2018
Db	1067	CATCTGGCTTGACGGGTTTCTTTGCCAGGACATGAACCTGCTGTCTGGCNAATGAACA	1126
QY	2019	CAGTGGGTGGGAAATAATGACTGCAGTGCAGATGAAGATGTTGGAGTGATGTTCTCTGA	2078
Db	1127	CCATGAATGGGAAAGACATTTTTCGAATCACAANTGAAGATGCTGGCGTGACATGTTCTGA	1186

[illegible]

Query Match		19.2%;	Score 886.8;	DB 22;	Length 4308;
Best Local Similarity		54.9%;	Pred. No. 1.8e-214;		
Matches 2061;		Conservative	0;	Mismatches 1527;	Indels 168; Gaps 9;
Qy	483	TCTGGTTTGGAGTGTGGATGGAACAACATCTCTGTTTCAGGAGAGAGTGGAGGTGAATTT	542		
Db	75	TCTGGAGCTGAGTGTGAAGGATGAGTCCATCGCTGTGAGGGAGAGTGGAAAGTGAAGCA	134		
Qy	543	CAAAGAAGGTGGGGACTATATCTGATGATGGTGGAACTTGATCTACTGTCGCGTGT	602		
Db	135	CAAAGGAGATGGGGCAGATGGATGGTTACAGTGGACATTTGAAGTATGATCTGTAGT	194		
Qy	603	GTGAGGCAACTAGGATGTCCATCTTCTTTTATTTCTTCTGGAGTGTGTTAATAGCCCTGC	662		
Db	195	GTGAGACAGCTGGGGTGTGAGCTGCCATTTGGTTTCTCTGGAGGGCTTATTTTGGGCC	254		
Qy	663	TGTATTGGCCCCATTTGGCTGGATGACATTTTATGCCCAGGGGAATGATTTGGCACTGTG	722		
Db	255	AGGACTTGGCCCCATTTGGCTTTGTATCTATCTGTGAAGGGACAGAGTCAACTGTTCAG	314		
Qy	723	GAATTGCAGACATCGTGTAGGGGAAATCATGCTGCTGCAATGAGGATGTCACATT	782		
Db	315	TGACTGTGAGCATCTAATATTAAGACTATCTGAATGATGGCTATATCATGTTGCGGA	374		
Qy	783	AACTTGTTATGATAGTAGTGTGTAAGTAAAGCTTGTAGTGGAACTAACCCTGTAT	842		
Db	375	TGCTGGAGTAGTCTGCTCAGGATTTG---TGCGTCTGGCTGGAGGGATGGACCCTGCTC	431		
Qy	843	GGGAGAGTAGCTGAAATCCAAAGGAGGTGGGGACCTGATGCCACCAATAGTGAA	902		
Db	432	AGGGCGAGTAGAAGTGCAATCTGAGAAAGCTTGATCCAGCTGCTGTGATGGGAATCTCAC	491		
Qy	903	CAATGCTCAGCTGATGCTGATGCAAGCAGTTGGGATGTGGAACCGCAGCTTCACTTCGC	962		
Db	492	ACTTGGCCACTGCCAGATCATCTGTGCAGAGTTGGTGTGGCAGGCTGTGTGTCCT	551		
Qy	963	TGGTGTGCTCATTTGACGTGAGGCTGTGATGTTGTATGGCTTGATGGTGTCTCTGCTC	1022		
Db	552	GGGACATGAGCTCTTCAGAGAGTCCAGTGCCAGGCTGCGGCTGAAGAGTTTCAGGTGTA	611		
Qy	1023	CGTTAATGAATCTTTCTTTGGGACTGCAGACATTCGCGAACCGTCAATTTTGACTGTCT	1082		
Db	612	GGGGAGGAGCTGAGCTCTGGGCTGCCCCAGAGTGCCCTGTCCAGGGGCGACGTGCA	671		
Qy	1083	TCATCAAAACCATGTGCTGTGATCTGCTCAGATGGAGCAGATTGGAACCTGCGACTAGC	1142		
Db	672	CCAGTGGATCTGCTCAGGTTGTTGTTTTCAGCATCTCAGA---AGTCCGGCTCATGAC	728		
Qy	1143	AGATGGAAGTAACAATTTTCAGGGAGAGTAGAGGTGAGAATTCATGAACAGTGGTGAC	1202		
Db	729	AAACGGCTCCTCTCAGTGTGAAGGCGAGGTGGAGATGAACATTTCTGGACAATGGAGGC	788		
Qy	1203	AATATGTACCAAGACTGGAAGATGAACAGCCCTTGTGTTGTGAAGCAGCTAGGATG	1262		
Db	789	GCTCTGTCCTCCCACTGGAGTGTGGCCAAATGCCAAATTTATCTCTGCTCAGCTCGGCTG	848		
Qy	1263	TCCGTTCAGCTCTTTGGCAGTCTGTGCTGCTGCTAAACCTTAGTAATGAAGCTAGAGCATTTG	1322		
Db	849	TGGAGTTGCCATCTCCACCCCGGAGGACCACACTTGGTGAAGAAGGTGATCATGCTT	908		
Qy	1323	GATAAACAGCATATCTTGGCACTGGGAATGAGTCAGCTCTCTGGGAGCTGCATATGATGG	1382		
Db	909	AACAGCCGATTTCACTGCTCTGGGCTGAGTCTCTCTGTTGGAGTTGTCTCTGTGACTGC	968		
Qy	1383	AAAAGCAAGCGACATGCTTCCGAAGATCATGATGCTGGAGTAATTTGTTCTGTAAAGGC	1442		
Db	969	CCTGGGTGGTCTGACTGTTCCCATGGCAACACAGCCCTCTGTGATCTGCTCAGGAAACCA	1028		
Qy	1443	AGATCTGG-----	1450		
Db	1029	GATCCAGGTGCTTCCCCAGTGCAACGACTCGGTGTCTCAACCTACAGGCTCTCGGGGCTC	1088		
Qy	1451	-----ACCTAAGGCTTGTGCGGGCTCATAG	1475		
Db	1089	AGAGACAGCGCCCTACTGCTCAGACAGCAGGCGAGCTCCGCTGTGGTGGAGCGGGCGG	1148		
Qy	1476	CCCTGTTATCGGAGATTTGGAGGTGAATACCAAGGAGAGTGGGGACTGTGTGTATGA	1535		
Db	1149	TCCCTGCCCGGGAGTGGAGATCCTTTGACAGAGGCTCTCTGGGCACCATCTGTGTATGA	1208		
Qy	1536	CAGATGGAGCAACAAGGAATGAGCTGTTGTGTAAACAATTTGGATGTGGAAGACCTAT	1595		
Db	1209	CGGCTGGACCTTGACGATGCCCGGTGTGTGAGGAGCTGGGCTGTGGAGAAGCCCT	1268		
Qy	1596	GCATGTTTGTGTATGACCTATTTTAAAGAACATCAGGACCTATTTTGGCTGGATGAGCT	1655		
Db	1289	CAATGCCACGGGCTGCTGCTCACTTCGGGGCAGGATCAGGCCCCATCTGTTGGACAACCT	1328		
Qy	1656	TCTTTGATTTGGAATGAGTCAATATCTGGGACTGTGAACACAGTGTGATGGGGAAGCA	1715		
Db	1329	CAACTGACAGGAAGGAGTCCCAGTGTGAGGTGSCCTTTCCCGGGGCTGGGGGCGACA	1388		
Qy	1716	TAATTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACATGGGGCT	1775		
Db	1389	CAACTGCAGACACAGCAGGACGCGGGGTCTATCTGCTCAGAGT-----TCCTGGCCCT	1442		
Qy	1776	GAGCTGTGGGGCGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGAGC	1835		
Db	1443	CAGGATGTGAGTGGAGGACAGCAGTGTCTGGTGGTGGAAAGTTTCTTACAATGGGAC	1502		
Qy	1836	CTGGGACAGCTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCA	1895		
Db	1503	CTGGGCGAGTGTCTGCCGTAAACCCCATGGAACATCACTGTGTCCACGATCTGCAGACA	1562		
Qy	1896	GCTGACGTGCCCATCTTCTATCATTTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGG	1955		
Db	1563	GCTTGGCTGTGGGACAGTGAACCCCTCACTCTCTTCTGTCTTAGAGAAGTTTAG	1622		
Qy	1956	AAAAATTTGGCTCGATGATGTTTCTGTGATGGAGATGAGTCACTCTGTCTGTATGAG	2015		
Db	1623	GCCACAGTGGTGGATAGATCCAGTGTGCGGAAACTGCACACTCTCTCTGSCAGTGCC	1682		
Qy	2016	GAACAGTGGTGGGGAATATGACTGCAGTGCACAGTGAAGATGTTGGAGTGTCTGTTTC	2075		
Db	1693	TCTGACCTTTGGAAATTAACAATCATGCTCTCCAAGGAGGAAGCCTATATCTGTGTG---	1739		
Qy	2076	TGATCATTCGGATATGAGGCTGAGGCTTGTGGTGTGGAAGCAGCAGCTGTGCTGGAAAAAT	2135		
Db	1740	---TGACACAGCAGACAGATCCGCTGTGGATGGAGTGTGCTCTCTTAGAGAAGTTAG	1796		
Qy	2136	TGAGTGAATGTCCAGGTGCCGTGGGAATTTGTGTGCTAATGCTGGGGGAATGAACAT	2195		
Db	1797	GGAGATCCTTGACCAAGGCTCTCTGGGCGACCATCTGTGATGACCGCTGGGACCTGGACGA	1856		
Qy	2196	TGCTCAAGTTGTTTCCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGACC	2255		
Db	1857	TGCCCGTGTGTGTGCAAGCAGCTGGGCTGTGGAGAAAGCCCTGGAGCCACTGTCTCTTC	1916		
Qy	2256	TCAATTTACAGAAAGAACATTACACATCTAATGTGCAATTTCTGCTGCACCTGGAGGGA	2315		
Db	1917	CTTCTTCGGGAGGATCAGGGCCCTCTGGCTGATGAGTGAAGTGAAGTGAAGGAGAGA	1976		
Qy	2316	AGCCTCTCTCTGGGATTTGTATACGATGGGAGTGGAAACAGACTGCGTGTCAATTAATAT	2375		
Db	1977	GTCCCAAGTATGGAGTGCCTTCTCGGGATGGGGGCAACACAACTGCAATCATCAAGA	2036		
Qy	2376	GGAGCAAGTTTGTCTCTCAGCCCCACAGGCCCCAGGCTGGTTGGAGCTGATATGCC	2435		
Db	2037	AGATGCAAGGATCTCTCTCAGGATTTGTGC-----GTCTGGCTGGAGGAGATGGACC	2090		
Qy	2436	CTGCTCTGAGCTGTGAGTGAACATGCACACATGCGGCTCTGCTGTGTGATTTCTGA	2495		
Db	2091	CTGCTCAGGGGAGTAGAAGTGCATTTGGAGAACGCTTGGACCCCGGAGTGTCTGTATGG	2150		
Qy	2496	TTTCTCTTCTCATGCTGCCAATGTCTGTGCAGAGAATTAATTTGAGAGATGCCATATC	2555		

```

Db 2151 CTTTACACCTCCACAGTGGCCAGTCTATCTGTGAGAGCTGGGATGTGGGAAGCTGTGC 2210
QY TCTTTCTGTGGGAGATCACTTTTGAAGGGAATGCTTAACCTTTGGGCCGCAAAAGTTCCA 2615
Db 2211 TGTCTTGGGACATGCCATTAGAGAGTCCGATGCCAGGTCTGGCTGAAGAGTTGAG 2270
QY GTGTGAAGGAGTGAACATCACTTGCATATATGCCCATTTGTCACATCCCGGAAGACAC 2675
Db 2271 GTGTGATGGGGGAGCCTGAGCTCTGGTCTGCCCCAGAGTCCCTGTCCAGGAGCCAC 2330
QY TTTCTATCCACAGCAGAGAGTGGAGTGTCTGTCTCCCGATATACAGATGTCCGACTTGT 2735
Db 2331 ATCTCTCCACAGTGGAGTGTCTGAGTGTCTGTCTGAGTGTACACAGAAATCCAGCTTAT 2390
QY G---AATGCAAAATCCAGTGTGACGGCAAGTGGAGATCAAGCTGCTTGGACATGGGG 2792
Db 2391 GAAAACGGCACCCTCAATGTGAGGGCAGTGGAGATGAAGATCTCTGGAGATGGAG 2450
QY CTCACTGTGTGACACCCACTGGGACCCAGAGATGCCCGTGTCTATGACAGAGCTGAG 2852
Db 2451 AGCGCTCTGTGCTCCACTGAGTGTGGCCAAATGCCAATGTGTCTGTCTGCTGAGCTCG 2510
QY CTGTGGGACTGTCTCTCAACACAGAGGAGAAATATATTGGAGAAAGTCTTCTGT 2912
Db 2511 CTGTGAGTGGCCATCTCCACCCAGAGGACCACACTTGGTGGAGGAGGTGTGATCAGAT 2570
QY GTGGGACACAGATTTTCAATGTAGGAAATGAGTCACTTCTGGATAAATCTCAAAATGAC 2972
Db 2571 CTCAACAGCCCAATTTCACTGTCTGAGGGCTGAGTCTCTCTGAGTGTCTCTGTGAC 2630
QY AGTCTTGTGAGCAGCTCCCTGTATCCATGGAATATCTGTCTGTATCTGTATCTGACAGAG 3032
Db 2631 TGCTTTGGGTGGGCTGACTGTTCCTATGTCACAGAGCTCTGTATCTGTCTGACGAAA 2690
QY CTGTACCCAGCCACTGTTCCTATGCTCCGCTCGCAATGTATCTGACCCATATTGTCTGCA 3092
Db 2691 CCACCCAGGTCTGCCCCAGTGCACAGCACTTCTGTCTCAACCTGACAGGTCTCGGC 2750
QY TCCAGAGGGAGTGTCTTGTATCTGTCTAGGAGCAAAAGCTCCGCTAGTGGATGGGA 3152
Db 2751 CTCAGAGGAGTGTCTCCCTACTGTCTCAGACAGCAGCAGCTCCGCTGTGTGACGGGG 2810
QY CAGCCGCTGTGCGGAGAGTAGAGATCTATCAGAGCCTCTCTGGGCAACCATCTGTGA 3212
Db 2811 CGTCTCCCTGCGGGGAGAGTGGAGATCTTGACAGGCTCTCTGGGCAACCATCTGTGA 2870
QY TGACGGCTGGGACCTGAGCGATCCCACTGTGTGTCTCAAAAGCTGGCTGTGGAGTGGC 3272
Db 2871 TGATGACTGGGACCTGGACGATCCCGTGTGTGTCTGACAGCAGCTGGCTGTGGAGAAGC 2930
QY CTTCAATGCCAGGTCTCTGCTCACTTTGGGAGGGGTGAGGGCCCATCTGGCTTGGATGA 3332
Db 2931 CCTCAATGCCAGGGTCTGCTCACTTGGGGGAGGATCAGGGCCCATCTGGCTGACGA 2990
QY CTTGAATGCCAGAGAGGAGTCCCATCTGTGGAGTGGCCCTTCCCGGGCTGGGGGCA 3392
Db 2991 CTTGAATGCCAGAGAGAGTCCCATCTGTGGAGTGGCCCTTCCCGGGCTGGGGGG 3050
QY GCACGACTGCAGGACCAAGGAGCAGCGGTCTCTCTCAGAGATTCACAGCCCTTGTAG 3452
Db 3051 GCACGACTGCAGACACAGGAGAGCAGCGGGTCTCTCTCAGAGTCTCTGCGCCCTCAG 3110
QY GCTCTACAGTGAACATGAACAGAGAGTGTGCTGGGAGATTTGGAAGTCTTCTATAACGG 3512
Db 3111 G-----ATGGTGAAGGAGGAGCAGCAGTGTGTGGTGGCTGTGGAGGTTTCTACAAAGC 3164
QY GACCTGGGCGAGCTGGGAGGAGGAACATCACCAGCCATGACAGGATGTGTGTCAG 3572
Db 3165 GACCTGGGCGAGTGTGCGGACAGCCCATGGAAGATATACATGTCTGCTGATCTGCGAG 3224
QY GCAGCTGGGCTGTGGGAGAGTGGAGTGTGTGACGCTTCGCCCTTTATCTAAGACAGGCTC 3632

```

```

Db 3225 ACAGCTTGGATGTGGGACAGTGGAAAGTCTCAACACCTCTGTGGTCTCAGGAAGGTTTC 3284
QY TGGTTTTCATGTGGGTGGATGACATTCAGTGTCTCTAAAGGCAATATCTCCATATGCGAGTG 3632
Db 3285 TAGACCCCGGTGGTAGATTAAATTCAGTGTGGAAATGGATACCTCTCTCTGCGAGTG 3344
QY CTTGTCTGCCATTTGGGAGGGAAGAATCTCCAGCCAGCAGAGAGACCTTGGATCAGATG 3752
Db 3345 TCTCTTGTGCCATGGAAATACAGTTTCATCTCTCCAAAGGAGAGCCCTACATCTCATG 3404
QY TGAAGATAGAATAA----- 3766
Db 3405 TGAAGGAAGAAGACCAAGAGCTGTCCAAGTGTCCGCTGTCACAGACAGAGAGAGCT 3464
QY --GAGTGCCTGGAGGAGACACCGAGTGTCTGGGAGAGTGGAGATCTGGCAGCAGCGCTC 3824
Db 3465 CCGCCTCAGGGGAGGAGACAGCGAGTGTCTCAGGCGGTGGAGGTGTGGCACAACGCTC 3524
QY CTGGGGCAGAGTGTGTGATGACTCTCTGGACCTTGGCCGAGCGGAGGTGTGTGTGAGCA 3884
Db 3525 CTGGGGCAGCGTGTGGGATGACTCTCTGGAGCCCTTGGCAGAGGCTGAGGTGTGTGTGAGCA 3584
QY GTTGGCTGTGGCTGTCTGCTGCTGCTGCTGAGGAGCGTCTGCTTGGCCAGGAACTGG 3944
Db 3585 GCTGGCTGTGGCCAGCGCTTGAAGCCGTGCGGTCTGACAGATTTGGCCCTGGAATGT 3644
QY AACCATCTGTTGGATGACATGCGGTGCAAGAGAAATAGTCAATTTCTATGGGACTGTCA 4004
Db 3645 GAGCATCTGGCTGGAGAGTGTGAGTGTGCGGGCGGGAGTCTCCCTGTGGGACTGTGT 3704
QY CGCCAAACCTTGGGACAGAGTGTGAGTGTGACACAGGAGAGAGTGTGGGTGAGGTGCTC 4064
Db 3705 TGCGGAGCCCTGGGCGCAGAGCAGTGTGCAAGCAGCAGGAGGATGCTGGTGTGAGGTGCTC 3764
QY TGACAGAGTGTGCTGAAATCACTGAATGCTCCTCCTCAGG 4100
Db 3765 TGGTGTGAAGACAAACATTGCCACGACACAGCAGG 3800

```

RESULT 12
AAAA3017
ID AAA43017 standard; cdNA; 820 BP.
XX
AC AAA43017;
XX
DT 21-AUG-2000 (first entry)
XX
Human secreted expressed sequence tag SEQ ID NO:1757.
DE
XX Human; mouse; xenopus; rat; secreted expressed sequence tag; EST;
KW expressed sequence tag; EST; probe; chemotactic; proliferative;
KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
KW antiviral; antidiabetic; antiasthmatic; vulnerary; antiparkinsonian;
KW antiulcer; osteopathic; neuroprotective; neurotropic; antipsoriatic;
KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;
KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KW central nervous system disorder; Alzheimer's disease; stroke;
KW Parkinson's disease; Huntington's disease; coagulation disorder;
KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
KW tumour; infection; depression; psoriasis; ss.
XX Homo sapiens..
XX
XX WO200021990-A1.
XX
PD 20-APR-2000.
XX
PF 15-OCT-1999; 99WO-US24205.
XX
XX 15-OCT-1998; 98US-0104435.
PR

D6 144 GTTTTAAATGGT

Qy 146 GTTTAAATGGAAACAGATTTCGAGGTTGAGGAGACGGTCCCTGCTCTGGGA 205

Db 144 GCTTTTAAATGGAAACAGATTTCGAGGTTGAGGAGACGGTCCCTGCTCTGGGA 203

Db 144 GTTTTAAATGGAACAGATTTCGAGTTGAGGCTGGTCAATGGAGACGGTCCCTGCTC

QY 206 CAGTGGAGGTGAATTCAGGACAGTGGGGACTGTGTGATGATGGGTGGAACACTA 265
|||||
Db 204 CAGTGGAGGTGAATTCAGGACAGTGGGGACTGTGTGATGATGGGTGGAACACTA 263
|||||
QY 266 CTGGCTCAACTGTCGTGTGCAACAGCTTGGATGCCA-TTTTCTTTGGCCATGTTTCGT 324
|||||
Db 264 CTGGCTCAACTGTCGTGTGCAACAGCTTGGATGCCA-TTTTCTTTGGCCATGTTTCGT 323
|||||
QY 325 TTTGGACAAGCCGTGACTAGACATGGAAATTTGGCTTGATGATGTTCTCTGTATGGA 384
|||||
Db 324 TTTGGACAAGCCGTGACTAGACATGGAAATTTGGCTTGATGATGTTCTCTGTATGGA 383
|||||
QY 385 AATGAGTCAGCTCTCTGGGAATGCAACACCGGAATGGGAAGCCATAACTGTTATCAT 444
|||||
Db 384 AATGAGTCAGCTCTCTGGGAATGCAACACCGGAATGGGAAGCCATAACTGTTATCAT 443
|||||
QY 445 GGAGAAGATGTTGGTGAACCTTTATGGTGAAGCCAACTCGGTTGAGGCTAGTGAT 504
|||||
Db 444 GGAGAAGATGTTGGTGAACCTTTATGGTGAAGCCAACTCGGTTGAGGCTAGTGAT 503
|||||
QY 505 GGAACAACTCTCTTTCAGGAGAGTGGAGTGAATTCRAAGAAAGTGGGGACTATA 564
|||||
Db 504 GGAACAACTCTCTTTCAGGAGAGTGGAGTGAATTCRAAGAAAGTGGGGACTATA 563
|||||
QY 565 TGTGATGATGGGTGGAACCTTCAATGCTGCTGCCGT 599
|||||
Db 564 TGTGATGATGGGTGGAACCTTCAATGCTGCTGCCGT 598
|||||

RESULT 15
AAH99326
ID AAH99326 standard; cDNA; 562 BP.
XX
AC AAH99326;
XX
DT 16-OCT-2001 (first entry)
XX
DE Human protein encoding cDNA sequence SEQ ID NO:161.
XX
KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiaggregant; haemostatic; vulnerary; antilulcer; osteopathic; eczema;
KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; neurotropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200153455-A2.
XX
PD 26-JUL-2001.
XX
PF 22-DEC-2000; 2000WO-US35017.
XX
PR 23-DEC-1999; 99US-0471275.
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457603/49.

P-PSDB; AAM25385.

Isolated human polynucleotides encoding polypeptides, useful for the treatment and diagnosis of e.g. cancer, ulcers and HIV infection

Claim 1; Page 369; 1217pp; English.

AAH99166 to AAH99904 encode the human proteins given in AAM25225 to AAM25963. The proteins can have activities based on the tissues and cells they are expressed in, such as: antiinflammatory; antirheumatic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant; central nervous system; virucide; anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; antiaggregant; antilulcer; osteopathic; vulnerary; antidiabetic; cytostatic; neuroprotective; antiallergic; antiasthmatic; antiparkinsonian; and immunostimulant. The proteins and polynucleotides encoding them can be used in gene therapy, antisense therapy and vaccine production. The proteins and polynucleotides are useful for screening for agonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmunity, genetic diseases, haematopoietic disorders, anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers, osteoporosis, severe combined immunodeficiency, eczema, allergic rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression, Alzheimer's disease, Parkinson's disease, neurodegenerative and neurological disorders.

Sequence 562 BP; 125 A; 131 C; 189 G; 117 T; 0 other;

Query Match 11.9%; Score 551.4; DB 22; Length 562;
Best Local Similarity 99.8%; Pred. No. 1e-129;
Matches 552; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3485 TGGGAGATTGGAAGTCTTCTATACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCA 3544
Db 10 CCGGAGATTGGAAGTCTTCTATACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCA 69
QY 3545 CCACAGCATAGCAGGCATTCTGTGCAGCAGCTGGGCTGTGGGAGAGTGGAGTTGCA 3604
Db 70 CCACAGCATAGCAGGCATTCTGTGCAGCAGCTGGGCTGTGGGAGAGTGGAGTTGCA 129
QY 3605 GCCTCGCCCTTTATCTAAGACAGCCTCTGTTTCATCTGGGTGGATGACATTCAGTGTG 3664
Db 130 GCCTCGCCCTTTATCTAAGACAGCCTCTGTTTCATCTGGGTGGATGACATTCAGTGTG 189
QY 3665 CTAAACCCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGCAGGCAAGATCTCCA 3724
Db 190 CTAAACCCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGCAGGCAAGATCTCCA 249
QY 3725 GCCCAGCAAGAGAGACCTGGATCAGATCTGAAGATAGATAAGAGTGGTGGAGGACA 3784
Db 250 GCCCAGCAAGAGAGACCTGGATCAGATCTGAAGATAGATAAGAGTGGTGGAGGACA 309
QY 3785 CCGAGTGTCTCTGGGAGAGTGGAGATCTGCCACGACAGGCTCTCTGGGGCCACAGTGTGTGATG 3844
Db 310 CCGAGTGTCTCTGGGAGAGTGGAGATCTGCCACGACAGGCTCTCTGGGGCCACAGTGTGTGATG 369
QY 3845 ACTCTGGGACCTGGCCGAGGCGGAAGTGTGTGTCACAGCTGGGCTGTGCTGTGCTC 3904
Db 370 ACTCTGGGACCTGGCCGAGGCGGAAGTGTGTGTCACAGCTGGGCTGTGCTGTGCTC 429
QY 3905 TGGCTGCCCTGAGGACGCTTCGTTTGGCCAGGGAACCTGGAACTCTGGTGGATGACA 3964
Db 430 TGGCTGCCCTGAGGACGCTTCGTTTGGCCAGGGAACCTGGAACTCTGGTGGATGACA 489
QY 3965 TCCGGTGC AAAAGGAATGAGTCATTTCTATGGGACTGTGACGCCCAACCCCTGGGACAGA 4024
Db 490 TCCGGTGC AAAAGGAATGAGTCATTTCTATGGGACTGTGACGCCCAACCCCTGGGACAGA 549
QY 4025 CTGACTGTGGACA 4037
|||||

Db 550 GTGACTGTGGACA 562

Search completed: May 12, 2003, 03:08:03
Job time : 683 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 03:08:10 ; Search time 50 Seconds
(without alignments)
3872.263 Million cell updates/sec

Title: US-09-759-130B-381
Perfect score: 8138
Sequence: 1 MMLPQNSWHIDFGRCCHQN.....CEDASDTSLGLVLPASEATK 1453

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_101002.*
- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
 - 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
 - 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
 - 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
 - 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
 - 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
 - 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
 - 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
 - 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
 - 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
 - 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
 - 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
 - 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
 - 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
 - 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
 - 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
 - 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
 - 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
 - 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
 - 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
 - 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
 - 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
 - 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. Is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8138	100.0	1453	22	Human TANGO 234.
2	7913	97.2	1413	22	Human TANGO 234 ma
3	7426	91.3	1319	22	Human TANGO 234 ex
4	6058	74.4	1120	22	Human secreted pro
5	4540	55.8	821	22	Human full-length
6	3441.5	42.3	1151	23	Human CD163 recept
7	3429	42.1	1121	22	Human polypeptide
8	3429	42.1	1124	22	Human polypeptide
9	3429	42.1	1124	22	Human polypeptide
10	3429	42.1	1156	23	Human CD163 recept

11	3424	42.1	1116	23	AAU97582	Human CD163 recept
12	3396.5	41.7	1149	23	AAU97583	Human CD163 recept
13	3396.5	41.7	1154	22	AAU97583	Human polypeptide
14	3082	37.9	1436	22	AAU97583	Bovine WCL protein
15	2435	29.9	1785	19	AAU97583	Human SRCR protein
16	1763.5	21.7	898	23	AAU09878	Novel human secret
17	1183	14.5	225	22	AAU25317	Human protein sequ
18	1048.5	12.9	1290	18	AAU07609	Rat von Ebner's gl
19	1033	12.7	186	22	AAU25385	Human protein sequ
20	971.5	11.9	552	22	AAU09447	Human sbg14862SPER
21	921	11.3	422	22	AAU09446	Human sbg14862SPER
22	822.5	10.1	822	20	AAU99087	Human serine prote
23	822.5	10.1	875	20	AAU99087	Human neurotrophin
24	753	9.3	143	22	AAU27881	Human contig polyp
25	725	8.9	141	22	ABU11125	Human M160 precurs
26	719.5	8.8	761	20	AAU99088	Mouse serine prote
27	719.5	8.8	761	20	AAU99088	Mouse neurotrophin
28	712.5	8.8	666	19	AAU64590	Human SRCR protein
29	691.5	8.5	347	19	AAU68200	Human scavenger re
30	691.5	8.5	347	20	AAU13369	Amino acid sequenc
31	691.5	8.5	347	22	AAU80237	Human PRO229 prote
32	691.5	8.5	347	23	ABU95434	Human angiogenesis
33	691.5	8.5	347	23	ABU84828	Human PRO229 prote
34	691.5	8.5	347	23	AAU83648	Human PRO protein,
35	690.5	8.5	347	19	AAU64537	Human liver cell c
36	634.5	7.8	757	21	AAU19127	Polypeptide isolat
37	634	7.8	757	21	AAU19127	Human Cys-rich sca
38	633.5	7.8	757	23	ABU07650	Human lysyl-oxidas
39	632.5	7.8	732	22	AAU11927	Human CG153 (Or C5
40	629	7.7	754	21	AAU00078	Murine lysyl oxida
41	627.5	7.7	753	22	AAU11935	Human CG153 (Or C5
42	625.5	7.7	743	23	AAU66060	Human lysyl oxidas
43	624.5	7.7	641	21	AAU12307	Human secreted pro
44	624.5	7.6	752	23	ABU07653	Human lysyl-oxidas
45	622.5	7.6	753	21	AAU00073	Human lysyl oxidas

ALIGNMENTS

RESULT 1

AAU66037
ID AAB66037 standard; Protein: 1453 AA.

XX AAB66037;

AC AAB66037;

DT 30-MAR-2001 (first entry)

XX Human TANGO 234.

DE TANGO protein; INTERCEPT protein; neurological disorder;

KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety.

OS Homo sapiens.

PN WO200077239-A2.

XX 21-DEC-2000.

XX 24-MAY-2000; 2000WO-US14858.

XX 14-JUN-1999; 99US-0333159.

XX (MILL-) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX WPI; 2001-032313/04.

DR N-PSDB; AAF45123, AAF45124.

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT screening assays and diagnostic assays and for the treatment of
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT disease -
XX
XX
XX Claim 8; Fig 2; 359pp; English.
XX
XX The present invention relates to TANGO or INTERCEPT proteins and coding
CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC sequences are useful for the treatment of neurological disorders such as
CC central nervous system (CNS) disorders, CNS-related disorders, focal
CC brain disorders, global-diffuse cerebral disorders and other
CC neurological and cerebrovascular disorders. The CNS disorders include
CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC autonomic function disorders such as hypertension and sleep disorders,
CC neuropsychiatric disorders, psychoactive substance use disorders,
CC anxiety, and bipolar affective disorder.
XX
XX Sequence 1453 AA;
SQ
Query Match 100.0%; Score 8138; DB 22; Length 1453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMLPONSWHIDGRCCCHONLFSAVVTCILLNSCLFSSNGTDLRLVNGDPCSGST 60
DB 1 MMLPONSWHIDGRCCCHONLFSAVVTCILLNSCLFSSNGTDLRLVNGDPCSGST 60
QY 61 VEVKQGGWGTVCDDGWNTTASTVYCKVKGCPFPFAMRFGQAVTRHGKIWLDDVSCYGN 120
DB 61 VEVKQGGWGTVCDDGWNTTASTVYCKVKGCPFPFAMRFGQAVTRHGKIWLDDVSCYGN 120
QY 121 ESALMEQHQREWSHNCYHGEDVGYNCYCEANLGLRLVDGNNCSGRVEVFKQERWGTIC 180
DB 121 ESALMEQHQREWSHNCYHGEDVGYNCYCEANLGLRLVDGNNCSGRVEVFKQERWGTIC 180
QY 181 DQGNLNTAAVYCRQLGCPSPFISGVVNSPAVLRPIWLDITLCOGNELALWNCRHGWG 240
DB 181 DQGNLNTAAVYCRQLGCPSPFISGVVNSPAVLRPIWLDITLCOGNELALWNCRHGWG 240
QY 241 NHDCSHNEDVTLCYDSSDLELRLVGGTNRNCRGVELKIQGRWGTVCVCHHKWNNAAADVVC 300
DB 241 NHDCSHNEDVTLCYDSSDLELRLVGGTNRNCRGVELKIQGRWGTVCVCHHKWNNAAADVVC 300
QY 301 KOLGCGTALHFAGLPHLOSGSDVWLDGVSCGNSFSLWDCRHSCTVNFDCVHNDVSVI 360
DB 301 KOLGCGTALHFAGLPHLOSGSDVWLDGVSCGNSFSLWDCRHSCTVNFDCVHNDVSVI 360
QY 361 CSDGADLELRLADGNNCSGRVEVRIHEQWWTICDQNKNEQALVVCQKCPFSVFGSR 420
DB 361 CSDGADLELRLADGNNCSGRVEVRIHEQWWTICDQNKNEQALVVCQKCPFSVFGSR 420
QY 421 RAKPSNEARDIWNISCTGNESALWDCYTDGKARTCFRRSDAGVICSADKADLRLVG 480
DB 421 RAKPSNEARDIWNISCTGNESALWDCYTDGKARTCFRRSDAGVICSADKADLRLVG 480
QY 481 AHSPCYGRLEVKYQGEWGTVCVCHDRSTNAAVVCQKCGKPMHVFGMTYFKAESGPIWL 540
DB 481 AHSPCYGRLEVKYQGEWGTVCVCHDRSTNAAVVCQKCGKPMHVFGMTYFKAESGPIWL 540
QY 541 DVVSCIGNESNIWDCSHSGWKHNCVHREDVITVCSGATWGLRLVGGNRCSGRLEVYF 600
DB 541 DVVSCIGNESNIWDCSHSGWKHNCVHREDVITVCSGATWGLRLVGGNRCSGRLEVYF 600
QY 601 QGRWGTVCDDGNNKAAAAYVVCQKCPSPSIICMGLGNASTGYGKIWLDDVSCOGDESILW 660
DB 601 QGRWGTVCDDGNNKAAAAYVVCQKCPSPSIICMGLGNASTGYGKIWLDDVSCOGDESILW 660
QY 661 SCRNNGWGNDCSHSDVGVICSDADMEELRLVGGSSRCAGKVEVNVQAVGILCANGWG 720
DB 661 SCRNNGWGNDCSHSDVGVICSDADMEELRLVGGSSRCAGKVEVNVQAVGILCANGWG 720

DB 661 SCRNNGWGNDCSHSDVGVICSDADMEELRLVGGSSRCAGKVEVNVQAVGILCANGWG 720
QY 721 MNIAEVVCRQLECCSATRVSRPEPHFTERTLHILMSNSCTGEASLWDCIRWENKOTACH 780
DB 721 MNIAEVVCRQLECCSATRVSRPEPHFTERTLHILMSNSCTGEASLWDCIRWENKOTACH 780
QY 781 LNMEASLICSAAHROPRLVADMPGSGRVEVYKHADTWRSVCDSDFSLSHAANVLCRELNGCD 840
DB 781 LNMEASLICSAAHROPRLVADMPGSGRVEVYKHADTWRSVCDSDFSLSHAANVLCRELNGCD 840
QY 841 AISLSVGDHFGKNGLTWAERFQCEGSETHALCPVIOHPEDTCHHSREVGVVCSRYTDV 900
DB 841 AISLSVGDHFGKNGLTWAERFQCEGSETHALCPVIOHPEDTCHHSREVGVVCSRYTDV 900
QY 901 RLNVKNSOCDGOVEINVLGHWSGLCDTHWDPEDARVLCQLSCGTALSTTGKYGIGERSV 960
DB 901 RLNVKNSOCDGOVEINVLGHWSGLCDTHWDPEDARVLCQLSCGTALSTTGKYGIGERSV 960
QY 961 RVWGRHFRHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLPCLANVSDPYLS 1020
DB 961 RVWGRHFRHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLPCLANVSDPYLS 1020
QY 1021 AVPEGSALICLEDKRLRLVDGSDRCAGRVEYHDFWGTICDDGWDLSDAHVVCKLGGC 1080
DB 1021 AVPEGSALICLEDKRLRLVDGSDRCAGRVEYHDFWGTICDDGWDLSDAHVVCKLGGC 1080
QY 1081 VAFNATVSAHFEGSGSPIWLDLNCGTGTHESLWQCPSRGWGQHDCHRHEDAGVICSEFTA 1140
DB 1081 VAFNATVSAHFEGSGSPIWLDLNCGTGTHESLWQCPSRGWGQHDCHRHEDAGVICSEFTA 1140
QY 1141 LRLYSETETESCAGRLEVYNGTWGSGRRNITTAIAGIVCRQLCGGNGVVSAPLSKT 1200
DB 1141 LRLYSETETESCAGRLEVYNGTWGSGRRNITTAIAGIVCRQLCGGNGVVSAPLSKT 1200
QY 1201 GSGFMWVDDIQCPTHISIWQCLSPWERRISSPAEETWITCEDRIRVRGGDECSGRVE 1260
DB 1201 GSGFMWVDDIQCPTHISIWQCLSPWERRISSPAEETWITCEDRIRVRGGDECSGRVE 1260
QY 1261 IWHAGSWGTVCDDSDWDLAEAEVVCQOLGCGSALAALRDASFQGTGTIWLDDMRCKGNS 1320
DB 1261 IWHAGSWGTVCDDSDWDLAEAEVVCQOLGCGSALAALRDASFQGTGTIWLDDMRCKGNS 1320
QY 1321 FLWDCHAPWGSDCHGHKEDAGVRCGSLKSLNASSGHLALILSSIFGLLLVLFILFL 1380
DB 1321 FLWDCHAPWGSDCHGHKEDAGVRCGSLKSLNASSGHLALILSSIFGLLLVLFILFL 1380
QY 1381 TWCVRQKOKHLPVSTRRRGSLLENLFHEMETCLKREDPHGTRTSDTTPNHGCCDASDT 1440
DB 1381 TWCVRQKOKHLPVSTRRRGSLLENLFHEMETCLKREDPHGTRTSDTTPNHGCCDASDT 1440
QY 1441 SLLGVLPASEATK 1453
DB 1441 SLLGVLPASEATK 1453
RESULT 2
AAB66039
ID AAB66039 standard; Protein; 1413 AA.
XX
XX AAB66039;
XX
XX 30-MAR-2001 (first entry)
XX
XX Human TANGO 234 mature protein.
XX
XX TANGO protein; INTERCEPT protein; neurological disorder;
KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety.
OS Homo sapiens.

XX WO200077239-A2.
 XX PD 21-DEC-2000.
 XX PF 24-MAY-2000; 2000WO-US14858.
 XX PR 14-JUN-1999; 99US-0333159.
 XX PA (MILL-) MILLENNIUM PHARM INC.
 XX PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
 XX DR WPI; 2001-032313/04.
 XX PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -
 XX Claim 8; Pages 276-281; 359pp; English.

XX The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.

XX Sequence 1413 AA;

Query Match 97.28; Score 7913; DB 22; Length 1413;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 FNGTDLRLVNGDPCSGTVEVFGQGWGTVCDGWNNTASTVYCKQLGCPFPFAMRFR 100
 Db 1 FNGTDLRLVNGDPCSGTVEVFGQGWGTVCDGWNNTASTVYCKQLGCPFPFAMRFR 60
 Qy 101 QGAVTRHGKIWDDVSCYGNESALWECOHREWSHNCYHGEDVGVNCGEANGLRLVDG 160
 Db 61 QGAVTRHGKIWDDVSCYGNESALWECOHREWSHNCYHGEDVGVNCGEANGLRLVDG 120
 Qy 161 NNSCGRVEVFKQERWGTICDDGNLNTAAVVCRLGCPSPFISGGVYNSPAVLRIWLD 220
 Db 121 NNSCGRVEVFKQERWGTICDDGNLNTAAVVCRLGCPSPFISGGVYNSPAVLRIWLD 180
 Qy 221 DILCOGNELALWNCRHRRGWHDCHSHNEDVLTCDYSSDLRLVGGTNRCMGRVELKIQ 280
 Db 181 DILCOGNELALWNCRHRRGWHDCHSHNEDVLTCDYSSDLRLVGGTNRCMGRVELKIQ 240
 Qy 281 GRWGTVCHHKWNNAAADVVCQLGCGTALHFAGLPHLOGSDVYVWLDGVCSCGNESFLWD 340
 Db 241 GRWGTVCHHKWNNAAADVVCQLGCGTALHFAGLPHLOGSDVYVWLDGVCSCGNESFLWD 300
 Qy 341 CRHSGTVNFCDLHQNDVSVICSDGADLELRADGSSNCGRVEVRIHQWWTICDONKN 400
 Db 301 CRHSGTVNFCDLHQNDVSVICSDGADLELRADGSSNCGRVEVRIHQWWTICDONKN 360
 Qy 401 EOALVVCQLGCPFVFGSRRAKPSNEARDIWNISICTGNESALWDCYDCKAKRTCFR 460
 Db 361 EOALVVCQLGCPFVFGSRRAKPSNEARDIWNISICTGNESALWDCYDCKAKRTCFR 420
 Qy 461 RSDAGVICSADKADLDRLVGAHSPCYGRLEVYKOGEWGTVCHDRWSTRNAAVVCKQLCG 520
 Db 421 RSDAGVICSADKADLDRLVGAHSPCYGRLEVYKOGEWGTVCHDRWSTRNAAVVCKQLCG 480

Qy 521 KPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCSEHSGWGHKNCVHREDVIVTCSGDAT 580
 Db 481 KPMHVFGMTYFKEASGPIWLDDVSCIGNESNIWDCSEHSGWGHKNCVHREDVIVTCSGDAT 540
 Qy 581 WGLRLVGGSNRCGRLEVYFQGRWGTVCDDGWNSSKAAAVVCSQDPCPSIIIGMGLNAST 640
 Db 541 WGLRLVGGSNRCGRLEVYFQGRWGTVCDDGWNSSKAAAVVCSQDPCPSIIIGMGLNAST 600
 Qy 641 GYKGIWLDVSCDGDSDLSWSCRNSGWNDCSHSDGVGICSDASDMELRVLGSSRCA 700
 Db 601 GYKGIWLDVSCDGDSDLSWSCRNSGWNDCSHSDGVGICSDASDMELRVLGSSRCA 660
 Qy 701 GKVEVNVQAGVILCANGWGMNIAEVVCRQLECGSAIRVSRPHFTERTLHLMNSGCT 760
 Db 661 GKVEVNVQAGVILCANGWGMNIAEVVCRQLECGSAIRVSRPHFTERTLHLMNSGCT 720
 Qy 761 GGEASLWDCIRWEWKQTACHLWMEASLICSASHRQPLRVGADMPGSGRVEVKHADTWRSVC 820
 Db 721 GGEASLWDCIRWEWKQTACHLWMEASLICSASHRQPLRVGADMPGSGRVEVKHADTWRSVC 780
 Qy 821 DSDFSLSHAANVLCRELNGDAISLSVGDHFGKNGLTWAEKFQCEGSETHLALCPTVQHP 880
 Db 781 DSDFSLSHAANVLCRELNGDAISLSVGDHFGKNGLTWAEKFQCEGSETHLALCPTVQHP 840
 Qy 881 EDTCIHSREVGVCVRYTDVRLVNGSKQCDGQVEINVLGHWSGLCDTHWDPEDARVLCRQ 940
 Db 841 EDTCIHSREVGVCVRYTDVRLVNGSKQCDGQVEINVLGHWSGLCDTHWDPEDARVLCRQ 900
 Qy 941 LSCGTALTSTTGKYGERSVRVWGHFRHFCNLGNESSLNDCQNTVLGAPPCIHGNTVSVICT 1000
 Db 901 LSCGTALTSTTGKYGERSVRVWGHFRHFCNLGNESSLNDCQNTVLGAPPCIHGNTVSVICT 960
 Qy 1001 GSLTOPLPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDFGWGTI 1060
 Db 961 GSLTOPLPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDFGWGTI 1020
 Qy 1061 CDDGWDLSDAHVVCKLGGVAFNATVSAHFEGEGSGPIWLDLNTCTGTESHLMQCPSRGW 1120
 Db 1021 CDDGWDLSDAHVVCKLGGVAFNATVSAHFEGEGSGPIWLDLNTCTGTESHLMQCPSRGW 1080
 Qy 1121 GOHDCRHKEDAGVICSEFTALRLYSETETESACGRLEVYNGTWGSGVRNLTITAGIV 1180
 Db 1081 GOHDCRHKEDAGVICSEFTALRLYSETETESACGRLEVYNGTWGSGVRNLTITAGIV 1140
 Qy 1181 CRQLCGGNGVYSLAPLSKTSGFMVDDIQCPTKTHISIWOCLSPAPWERRISSPAEETWI 1240
 Db 1141 CRQLCGGNGVYSLAPLSKTSGFMVDDIQCPTKTHISIWOCLSPAPWERRISSPAEETWI 1200
 Qy 1241 TCEDRIRVRGGTECSGRVEIWHAGSWGTVCCDSDWDLAEAEVVCQQLGCGSALAALRDAS 1300
 Db 1201 TCEDRIRVRGGTECSGRVEIWHAGSWGTVCCDSDWDLAEAEVVCQQLGCGSALAALRDAS 1260
 Qy 1301 FQOGTGTIWLDDMRCKGNESFLWDCHAKPWGSDCGCHKEDAGVRCSSGSLKSLNASSGHL 1360
 Db 1261 FQOGTGTIWLDDMRCKGNESFLWDCHAKPWGSDCGCHKEDAGVRCSSGSLKSLNASSGHL 1320
 Qy 1361 ALILSSIFGLLLVLFLFTWCRVOKKHLPLRYSTRRRGSLLENLPHMETCTCLKREDP 1420
 Db 1321 ALILSSIFGLLLVLFLFTWCRVOKKHLPLRYSTRRRGSLLENLPHMETCTCLKREDP 1380
 Qy 1421 HGTRTSDTTPNHGCEADSDTSLGLVLPASEATK 1453
 Db 1381 HGTRTSDTTPNHGCEADSDTSLGLVLPASEATK 1413

RESULT 3
 AAB66040
 ID AAB66040 standard; Protein; 1319 AA.
 XX AAB66040;
 AC AAB66040;
 XX
 DT 30-MAR-2001 (first entry)
 XX

Human TANGO 234 extracellular domain.

TANGO protein; INTERCEPT protein; neurological disorder; central nervous system; focal brain disorder; bipolar affective disorder; global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease; senile dementia; Huntington's disease; amyotrophic lateral sclerosis; parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep; neuropsychiatric; psychoactive substance use; anxiety.

Homo sapiens.

WO200077239-A2.

21-DEC-2000.

24-MAY-2000; 2000WO-US14858.

14-JUN-1999; 99US-0333159.

(MILL-) MILLENNIUM PHARM INC.

McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

WPI: 2001-032313/04.

TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for screening assays and diagnostic assays and for the treatment of neurological diseases such as Alzheimer's, Parkinson's and Huntington's disease.

Claim 8; Pages 281-287; 359pp; English.

The present invention relates to TANGO or INTERCEPT proteins and coding sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057, AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding sequences are useful for the treatment of neurological disorders such as central nervous system (CNS) disorders, CNS-related disorders, focal brain disorders, global-diffuse cerebral disorders and other neurological and cerebrovascular disorders. The CNS disorders include Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic lateral sclerosis, parkinson's, Gilles de la Tourette's syndrome, autonomic function disorders such as hypertension and sleep disorders, neuropsychiatric disorders, psychoactive substance use disorders, anxiety, and bipolar affective disorder.

Sequence 1319 AA;

Query Match 91.3%; Score 7426; DB 22; Length 1319;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 FNGTDLRLVNGDPCSGTVEYKFGQGTVCDDGNTTASTVVCQKLGCPSPFAMFRF 100
 Db 1 FNGTDLRLVNGDPCSGTVEYKFGQGTVCDDGNTTASTVVCQKLGCPSPFAMFRF 60

QY 101 GOAVTRHGKIWLDDVSCYGNESALWECQREHWSHNCYHGEDVGVNVCYGEANLGLRLVDG 160
 Db 61 GOAVTRHGKIWLDDVSCYGNESALWECQREHWSHNCYHGEDVGVNVCYGEANLGLRLVDG 120

QY 161 NNSCSGRVEVKFOERWGTICDDGNNLNTAAVVCROLGCPSPSFSSGVVNSPAVLRPIWLD 220
 Db 121 NNSCSGRVEVKFOERWGTICDDGNNLNTAAVVCROLGCPSPSFSSGVVNSPAVLRPIWLD 180

QY 221 DILCOGNELALWNCRRHGWGNDCHSHNEDVTLTCYDSSDLELRVGGTNRGMRVRLKIQ 280
 Db 181 DILCOGNELALWNCRRHGWGNDCHSHNEDVTLTCYDSSDLELRVGGTNRGMRVRLKIQ 240

QY 281 GRGTVCHHKNNAADVVCQKLGCGTALHFAGLPHLQSGSDVWLDGVSCSNESFLWD 340
 Db 241 GRGTVCHHKNNAADVVCQKLGCGTALHFAGLPHLQSGSDVWLDGVSCSNESFLWD 300

QY 341 CRHSGTVNFDCFLHONDVSVICSDGADLELRADGNNCSGRVEVRIHEQWWTICDQWKN 400
 Db 341 CRHSGTVNFDCFLHONDVSVICSDGADLELRADGNNCSGRVEVRIHEQWWTICDQWKN 360

RESULT 4
 AA000396
 ID AA000396 standard; Protein; 1120 AA.
 XX
 AC
 AA000396;
 XX

DT 04-JUL-2001 (first entry)
XX Human secreted protein, POLY8.
XX Human secreted protein; therapeutic; diagnostic; human; cancer.
XX Homo sapiens.
XX WO200119856-A2.
XX 22-MAR-2001.
XX 13-SEP-2000; 2000WO-US25106.
XX 13-SEP-1999; 99US-0153629.
XX 16-SEP-1999; 99US-0154520.
XX 20-SEP-1999; 99US-0154762.
XX 13-OCT-1999; 99US-0159231.
XX 12-SEP-2000; 2000US-0659634.
XX (CURA-) CURAGEN CORP.
XX Shimkets RA, Fernandes E, Herrmann JL, Liu X, Yang M, Boldog FL;
PI WPI; 2001-244781/25.
XX N-PSDB; AAS01217.
XX New POLYX polypeptide useful for treating or preventing a POLYX
PT associated disorder, e.g. cancer -
XX Claim 9; Page 25-29; 152pp; English.
XX The sequence represents the amino acid sequence of human secreted
CC protein, POLY8. POLYX nucleic acids, polypeptides and antibodies to
CC POLYX can be used for treating or preventing a POLYX associated disorder
CC in a subject, preferably a human. These can be used in the manufacture of
CC a medicament for treating a syndrome associated with a human disease
CC selected from a POLYX-associated disorder, where the therapeutic is a
CC POLYX polypeptide, a POLYX nucleotide or a POLYX antibody. They may also
CC be used to screen for a modulator of activity, or latency, or
CC predisposition to a POLYX associated disorder, e.g. cancer.
XX
SQ Sequence 1120 AA;
Query Match 74.4%; Score 6058; DB 22; Length 1120;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1088; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 363 DGADLELRADGNSNCGRVEYRIHEQWMTICDQNKNEQALVVCQKQCPFSVFGSRR 422
Db 30 DGADLELRADGNSNCGRVEYRIHEQWMTICDQNKNEQALVVCQKQCPFSVFGSRR 89
QY 423 KPSNEARDIWNISICTGNESALWDCTYDGAQKRTCFRRSDAGVICSOKADLDLRLVGAH 482
Db 90 KPSNEARDIWNISICTGNESALWDCTYDGAQKRTCFRRSDAGVICSOKADLDLRLVGAH 149
QY 483 SPYGRLEVKYOGEGTGTCHDRWTRNAAVVCQKQCPKPMHVFGMTYFKEASGPIWLD 542
Db 150 SPYGRLEVKYOGEGTGTCHDRWTRNAAVVCQKQCPKPMHVFGMTYFKEASGPIWLD 209
QY 543 VSCIGNESINWDCSHSGKHKNCVHREDIVITCSGDATWGLRLVGGSNRCSGRLEYIFOG 602
Db 210 VSCIGNESINWDCSHSGKHKNCVHREDIVITCSGDATWGLRLVGGSNRCSGRLEYIFOG 269
QY 603 RGTGTVDDGWSNKAAYVVCQKQCPSSIGMGLNASTGYGKIWLDVSCDGEDSLWSC 662
Db 270 RGTGTVDDGWSNKAAYVVCQKQCPSSIGMGLNASTGYGKIWLDVSCDGEDSLWSC 329
QY 663 RNSGWNDCSHSDVGVTCSDASPMELRVGSSRCAGKVEVNVQAGVILCANGWGN 722
Db 330 RNSGWNDCSHSDVGVTCSDASPMELRVGSSRCAGKVEVNVQAGVILCANGWGN 389
QY 723 TAEVVCRLQECGSAIRVSRPHFTERTLHILMSNSGCTGGEASLWDCIRWENKQTACHLN 782

Db 390 TAEVVCRLQECGSAIRVSRPHFTERTLHILMSNSGCTGGEASLWDCIRWENKQTACHLN 449
QY 783 MEASLICSAPROPLRVGADMPGSRVEVKHATWRSVCDSDFSLAHVLCRELNGDAI 842
Db 450 MEASLICSAPROPLRVGADMPGSRVEVKHATWRSVCDSDFSLAHVLCRELNGDAI 509
QY 843 SLVSGDHFQKGNGLTWAERFQCEGSETHLALCPIVQHPEDTCTIHSREVGVVCSRYTDVRL 902
Db 510 SLVSGDHFQKGNGLTWAERFQCEGSETHLALCPIVQHPEDTCTIHSREVGVVCSRYTDVRL 569
QY 903 VNGKSQCDGQVEINVLGHWSLCLDTHWDPEDARVLCRLQSCGTALSTTGKYGIGERSVRV 962
Db 570 VNGKSQCDGQVEINVLGHWSLCLDTHWDPEDARVLCRLQSCGTALSTTGKYGIGERSVRV 629
QY 963 WGRHFRHCLGNESLLDNCOMTVLGAPCIHGNVTSVLCSTSLTOPLPCLANVSDPYLSAV 1022
Db 630 WGRHFRHCLGNESLLDNCOMTVLGAPCIHGNVTSVLCSTSLTOPLPCLANVSDPYLSAV 689
QY 1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDFWGTICDDGWDLSDAHVVCQKLGCGVA 1082
Db 690 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDFWGTICDDGWDLSDAHVVCQKLGCGVA 749
QY 1083 FNATVSAHFEGSGPIWLDLNLCTGTGESHLMQCPSRGWGQDCRHKEDAGVICSEFTALR 1142
Db 750 FNATVSAHFEGSGPIWLDLNLCTGTGESHLMQCPSRGWGQDCRHKEDAGVICSEFTALR 809
QY 1143 LYSETETESCAGRLEVFYNGTWGSGVRRNITTAIAGIVCRLCGGNGVVS LAPLSTGTS 1202
Db 810 LYSETETESCAGRLEVFYNGTWGSGVRRNITTAIAGIVCRLCGGNGVVS LAPLSTGTS 869
QY 1203 GFWMVDDIOCPKTHISIWQCLSPAPERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 1262
Db 870 GFWMVDDIOCPKTHISIWQCLSPAPERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 929
QY 1263 HAGSMGTVCDDSDWDLAEAEVVCQKQCGSALALRDASFQGGTGTIWLDDMRCKGNESFL 1322
Db 930 HAGSMGTVCDDSDWDLAEAEVVCQKQCGSALALRDASFQGGTGTIWLDDMRCKGNESFL 989
QY 1323 WDCHAKPMQSGDCGHKEDAGVRCGOSLKSALNASSGHLALILSSIFGLLLVLFILFTW 1382
Db 990 WDCHAKPMQSGDCGHKEDAGVRCGOSLKSALNASSGHLALILSSIFGLLLVLFILFTW 1049
QY 1383 CRVQKOKHPLRVSTRRRGSLLENLFHEMETCLKRDPHGTRTSDTTPNHGCCDASDTSL 1442
Db 1050 CRVQKOKHPLRVSTRRRGSLLENLFHEMETCLKRDPHGTRTSDTTPNHGCCDASDTSL 1109
QY 1443 LGVLPASEATK 1453
Db 1110 LGVLPASEATK 1120
RESULT 5
AAU27709
ID AAU27709 standard; Protein; 821 AA.
XX
XX AAU27709;
XX 18-DEC-2001 (first entry)
XX Human full-length polypeptide sequence #34.
XX Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
KW nervous system disorder; inflammatory disorder; cell differentiation;
KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
KW cytostatic; antirheumatic; antiarthritic; vulnery; antinflammatory;
KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW neuroprotective; osteopathic; antidiabetic; antiallergic;
KW immunostimulant; analgesic; gene therapy.
XX

OS Homo sapiens.
 PN WO200164834-A2.
 XX
 XX
 PD 07-SEP-2001.
 XX
 XX 26-FEB-2001; 2001WO-US04926.
 XX
 XX 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 PR 17-JUN-2000; 2000US-0597707.
 PR 14-JUL-2000; 2000US-0616807.
 PR 19-SEP-2000; 2000US-0664641.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
 PI Drmanac R;
 DR WPI; 2001-589862/66.
 DR N-PSDB; AAS44609.
 XX
 XX Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis, treatment of
 PT cancer, neurological, inflammatory disorders and for use in arrays for
 PT detection
 PT
 PS Claim 10; SEQ ID No 206; 153pp; English.
 XX
 XX Sequences AAU27676-AAU28019 represent full-length polypeptides and
 CC contig polypeptides of the invention. The proteins and their associated
 CC DNA sequences are useful for the treatment, diagnosis and prevention of
 CC various types of disorder in a mammalian subject such as a human, dog,
 CC monkey, mouse, hamster or rat. The disorders include cancers such as
 CC leukemia, lymphoma and neuroblastoma, autoimmune disorders such as
 CC multiple sclerosis, connective tissue disease, rheumatoid arthritis,
 CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
 CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
 CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
 CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
 CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
 CC bowel disease. The sequences exhibit activity relating to angiogenesis,
 CC cell proliferation, cell differentiation, stem cell growth factor,
 CC activin or inhibin. Therefore, they can be used to manipulate stem cells
 CC in culture to give rise to neuroepithelial cells that can be used to
 CC augment or replace cells damaged by illness, accidental damage or genetic
 CC disorders. The sequences may also be used for regeneration of bone
 CC cartilage, tendons and ligaments and in tissue repair and burn healing.
 CC Note: Some sequences for this patent did not form part of the printed
 CC specification, but were obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 821 AA;
 SQ
 Query Match 55.8%; Score 4540; DB 22; Length 821;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 820; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 633 MGLGNASTGKTKWLLDDVSCDGEDSLWSCRNSGWNNDCHSHEDVGVICSDASDMELRL 692
 DB 1 MGLGNASTGKTKWLLDDVSCDGEDSLWSCRNSGWNNDCHSHEDVGVICSDASDMELRL 60
 QY 693 VGGSSRCAGRVENVQAVGILCANGWMNIAEVVCRQLECGSAIRVSRPFHETRLHI 752
 DB 61 VGGSSRCAGRVENVQAVGILCANGWMNIAEVVCRQLECGSAIRVSRPFHETRLHI 120
 QY 753 LMSNSGCTGGEASLWDCIRWEKQOTACHLNNEASLCSAHRQRLVGMADMPGSRVEVKH 812
 DB 121 LMSNSGCTGGEASLWDCIRWEKQOTACHLNNEASLCSAHRQRLVGMADMPGSRVEVKH 180
 QY 813 ADTWRSVCDSDFSLHAANVLCRELNCDAISLSVGDHFGKNGLTWAEKFCQEGSETHLA 872
 DB 181 ADTWRSVCDSDFSLHAANVLCRELNCDAISLSVGDHFGKNGLTWAEKFCQEGSETHLA 240
 QY 873 LCPIVQHPEDTCTHSREVGVCSTRYTDVRLVNGKSCQDQVEINVLGHWSLCTDTHWDPE 932
 DB 241 LCPIVQHPEDTCTHSREVGVCSTRYTDVRLVNGKSCQDQVEINVLGHWSLCTDTHWDPE 300
 QY 933 DARVLCRQLSCGTALTSTTGKYGIGERSVRVWGHRRFCHLCNGLNESLLDNCQMTVLGAPPCIHG 992
 DB 301 DARVLCRQLSCGTALTSTTGKYGIGERSVRVWGHRRFCHLCNGLNESLLDNCQMTVLGAPPCIHG 360
 QY 993 NTVSVICTGSLTQPLPCLANVSDPYLSAVPEGSALICLEDKRLRLVGDGSRGAGRVEIY 1052
 DB 361 NTVSVICTGSLTQPLPCLANVSDPYLSAVPEGSALICLEDKRLRLVGDGSRGAGRVEIY 420
 QY 1053 HDGFWCTICDDGWDLSDAHVVCKLGGVAFNATVSAHFEGSGSPIWLLDNLCTGTESH 1112
 DB 421 HDGFWCTICDDGWDLSDAHVVCKLGGVAFNATVSAHFEGSGSPIWLLDNLCTGTESH 480
 QY 1113 WQCPSPRGWGHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSGVRRI 1172
 DB 481 WQCPSPRGWGHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSGVRRI 540
 QY 1173 TTAAGTIVCRQLCGGNGVVS LAPLSKTGSGFMWDDIOCPKTHISIWOCLSPAPWERRIS 1232
 DB 541 TTAAGTIVCRQLCGGNGVVS LAPLSKTGSGFMWDDIOCPKTHISIWOCLSPAPWERRIS 600
 QY 1233 SPAEETWITCEDIRVRGDTGTECSGRVEIWHAGSWGTVCDSDWDLAEAEVVCQOLGCGSA 1292
 DB 601 SPAEETWITCEDIRVRGDTGTECSGRVEIWHAGSWGTVCDSDWDLAEAEVVCQOLGCGSA 660
 QY 1293 LAALRASFGQGTITWLLDMRCKGNESFLWDC HAKPWGSDCGCHKEDAGVRCGSGSLKS 1352
 DB 661 LAALRASFGQGTITWLLDMRCKGNESFLWDC HAKPWGSDCGCHKEDAGVRCGSGSLKS 720
 QY 1353 LNASSGHLALILSSIFGLLLVFLFTWCRVQKQKHLPLRVSTRRRGSLLEENLFHME 1412
 DB 721 LNASSGHLALILSSIFGLLLVFLFTWCRVQKQKHLPLRVSTRRRGSLLEENLFHME 780
 QY 1413 TCLKREDPHGTRTSDDTPNHGCEDASDTSLGLVLPASEATK 1453
 DB 781 TCLKREDPHGTRTSDDTPNHGCEDASDTSLGLVLPASEATK 821
 RESULT 6
 AAU97585
 ID AAU97585 standard; Protein; 1151 AA.
 XX
 XX AAU97585;
 AC
 XX AAU97585;
 DT 27-AUG-2002 (first entry)
 XX
 XX Human CD163 receptor extracellular variant protein sequence.
 DE
 XX Human; CD163; receptor; haptoglobin-haemoglobin complex; Hp-Hb complex;
 KW CD163 receptor variant; antibiotic; anticancer drug; anti-HIV drug;
 KW human immunodeficiency virus; haemolysis; haematological condition;
 KW aplastic anaemia; iron-deficiency anaemia; megaloblastic anaemia;
 KW sickle-cell anaemia; polycythemia; malaria; leukaemia; myelodysplasia;
 KW lymphoma; leucopenia; splenectomy; inflammation; infection; cancer;
 KW autoimmunity; immunodeficiency; CD163 receptor extracellular variant.
 XX
 OS Homo sapiens.
 XX
 XX WO200232941-A2.
 PN
 XX
 PD 25-APR-2002.
 XX
 XX 12-OCT-2001; 2001WO-DK00671.
 PF
 XX 16-OCT-2000; 2000DK-0001543.
 PR
 XX 11-JAN-2001; 2001DK-0000039.
 PR
 XX 22-FEB-2001; 2001US-270120P.
 PR
 XX

(PROT-) PROTEOPHARMA APS.
 Moestrup S, Moller HJ;
 WPI: 2002-452380/48.
 New haptoglobin-haemoglobin complex or its mimics or receptors, useful in gene therapy, particularly for treating an individual suffering from e.g. haemolysis, anaemia, inflammation, infection, cancer or autoimmunity

Claim 1: Fig 5; 70pp; English.

The present invention relates to a new haptoglobin-haemoglobin (Hp-Hb) complex, or its part or mimic, being operably linked to a substance, capable of binding a CD163 receptor and/or a CD163 receptor variant. The Hp-Hb complex or its mimic, the CD163 variant or the CD163 receptor are useful as a medicament or in the manufacture of a medicament for treating an individual. The medicament comprises at least one drug to be delivered to a cell expressing a CD163 receptor or a CD163 variant. The medicament may also comprise at least one gene to be delivered to a cell expressing a CD163 receptor or a CD163 variant. The medicament is an antibody or an anticancer drug, e.g. an anti-HIV (human immunodeficiency virus) drug. In particular, the Hp-Hb complex or its mimic, the CD163 variant or the CD163 receptor, or the medicament are useful for treating haemolysis and/or other haematological conditions, aplastic anaemia, iron-deficiency anaemia, megakaryoblastic anaemia, sickle-cell anaemia, polycythemia, malaria, leukaemia, myelodysplasia, lymphoma, leucopenia, splenectomy, inflammation, infection, cancer, autoimmunity or immunodeficiency. The method involves drug-delivery treatment or gene delivery treatment of an individual. The medicament is also useful for inhibiting (in vivo) uptake of haemoglobin in cells presenting a CD163 receptor. The CD163 variant is also useful in diagnosing, monitoring and/or controlling a condition in an individual. In particular, the CD163 variant is useful for identifying macrophages in a biological sample from an individual, for identifying at least one Hp-Hb complex in serum and/or plasma of an individual, or for removing at least one Hp-Hb complex in serum and/or plasma of an individual. The present amino acid sequence represents the human CD163 receptor extracellular variant of the invention.

Query Match 42.3%; Score 3441.5; DB 23; Length 1151;
 Best Local Similarity 55.2%; Pred. No. 2.7e-239;
 Matches 608; Conservative 166; Mismatches 311; Indels 17; Gaps 6;

364 GADLELRADGNSNCSGRVEVRIHQWTTCDQNKNEQALVCKQLGCPFSVFGSRRAK 423
 41 GTDKELRLVDGKNCGRVEVKVQEEWGTVCNNGWSMEAVSVICNLGCPATAKPGWAN 100
 424 PSNEARDIWNISCTGNSALWDCTYDGRKRT-CFRSDAGVICSADKADLDRLVGAH 482
 101 SSAGSGRIEMHWDVSCRGNSALMDCKDHGKHSNCTHOODAGVTCSDGNSLEMLRTGG 160
 483 SPYCYRLVRYQGEWGVFVCHDRSTNAAYVCKQLGCGKPMHVFMTYFKAEAGPIWLD 542
 161 NMCSGRIEIKFQWGTVCDDNFNIDIASVICRQLECGSAVSFGSSNFGEGSGPIWFD 220
 543 VSCIGNESNTWDCHEGSHGKNCVHREDIVTCSGDATWGLRVGGNSGRLEVFYFQ 602
 221 LICNGNESALWNCQKQWGHKNCDAEDAGVICSAGKADLSRLRLVDGVTECSGRLEVF 280
 603 RWGTCVDDGNSKAAAVVCSQLDPCPSIIIGMGLNASTGVKILWDVSCDGEDSLWSC 662
 281 EMGTICDDGNSDYDAAVACKOLGCPATAVTAIGRVNASKGFHILWDSVSCGHEPAWQC 340
 663 RNSGWNDCSHSDVGVICSADSMELRLVGGSSRCAGKVEVNVQAGVILCANGWMN 722
 341 KHEWGHYCNHNEADAGVTCSDGSDLELRLLGGSGRCAGTVEIORLLKVCVDRGHLK 400
 723 IAEVVCQLGECGSAIRVSRPHFTERTLHILMSNSGCTGGEASLWDCIRWENKOTACHLN 782

401 EADVVCRLCGGSALKTSYQVYSKIQATNTWTLFSSCNGNETSLWDCKNQWGLTCDHY 460
 783 MEASLCSAHRQRLVGADMPGSRVEVKHADTWRSVCDSDFSLHAANVLCRELNCDAI 842
 461 EEAKITCSAHRPRLVGDDIPGSRVEVKHGDWTGSGICSDSDFSLEAASVLCRELQCTTV 520
 843 SLVSGDHFHKGNGLTWAEKFCGSETHLALCPVQHPEDTCIHSREVGVCVSRVTVRL 902
 521 SIIGAHFNGGNGQIWAEEFCQEGHESHLCLCPVAPREGTCSHSRDVGVCVSRVTEIRL 580
 903 VNGKSQCDGQVEINVLGHWSGLCDTHWDPEDARVLCRLQSCGTALSTTGKYGIGERSVVR 962
 581 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLCGVALSTPGGARFCGNGQI 640
 963 WGRHFHCLNGLDNCQMTVLGAPPCIHGNTVSVITGSLTQPLFCLANVPDYLISAV 1022
 641 WRHMFHCTGTGTHMGDCPVLTALGASLCPSEQVAVSICSGNQSLTSSCNSGLPRTPT 700
 1023 PEGSALICLEDKRLRLVDGSRGACGRVEIYHDGFWGTICDDGWDLSDAHVVCOKLGGVA 1082
 701 PERSAVACIESGQLRVNGGRCAGRVEIYHEGSGWTICDDSWDLSDAHVVCRLGCGEA 760
 1083 FNATVSAHFEGEGSGPIWLDLNLCTGTESHLWQCPGSRGQHDCHRHEDAGVICSEFTALR 1142
 761 INATGSAHFEGEGTGTWLDKMKCKNGKESRIWQCHSHGQNGQNCRHEDAGVICSEFNLSR 820
 1143 LYSETETESACGRLEVPYNGTWGVSGRNITTAITAGIVCRLQCGGNGVVS LAPLSKTS 1202
 821 LTSEASREACAGRLEVPYNGAWGTGKSSMETTVGVVVCRLQCADKGINPASLDKMS 880
 1203 GFMYDDIOCPKTHISTWQCLSPAWERRISSPABETWITCEDRIRVRGGDTECSGRVEIW 1262
 881 IPMVDNVQCPKGPDTLWQCPSPWEXRLASPEETWITCDNKIRLOEGPTSCSGRVEIW 940
 1263 HAGSWGTVCDSDWDLAEVVCVQQLGCGSALALRADSFQGTGTIWLDDMRCKGNSFEL 1322
 941 HGGSWGTVCDSDWDLDAQVVCVQQLGCGPALKAFKAEFGQGTGPIWLNEVCKAGNESSL 1000
 1323 WDHAKPWGSDCGHKEDAGVRCGSGSLK-----SLNASSGHLALILSSIFGLLLVLF 1376
 1001 WDCPARWGHSECHGEDAAVNCITDISVQKTPQKATGRSSRSQSFATVGLGVVLLAIF 1060
 1377 I-LFLTWCRVQKOKHLRLRVSTRRRGSLLENLFHEMETCLKREDPHGTSTDDTPNHG 1434
 1061 VALFFLTKKRRQRQ---LAVSSRGENLVHQIQYRENNNSCLNADDLDLMSSEN--SHES 1115
 1435 EDASDTSLLGV---LPASEATK 1453
 1116 ADFSAAELISVSKFLPIISGMEK 1137

RESULT 7
 AAK39493
 ID AAK39493 standard; Protein; 1121 AA.
 XX
 AC AAK39493;
 XX
 AC
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 2638.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW ankyrotrophic lateral sclerosis; Shy-drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.
 XX 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR N-PSDB; AA158649.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PT
 XX Example 4; SEQ ID NO 2638; 10078pp; English.
 XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA38642-AA42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX Sequence 1121 AA;
 SQ

Query Match 42.1%; Score 3429; DB 22; Length 1121;
 Best Local Similarity 56.2%; Pred. No. 2.1e-238;
 Matches 598; Conservative 161; Mismatches 294; Indels 12; Gaps 4;

QY 364 GADLELRADGSGNCGRVEVRIHROWTICDONKNEQALVVCQKOLGCPFVFGSRRAK 423
 DB 46 GTDKELRLVDGENKCSGRVEVKEVQEEWGTVCNNGWSMEAVSVCNQLGCPPTAIAKPGWAN 105
 QY 424 PSNEARDIWINISCTGNESALWDCTYDCKAKRT-CFRRSADAGVCSDKADLDLRLVGAH 482
 DB 106 SAGSGRIWMDHVSRCNSESALWDCKHDGWHKSNCTHOODAGVTCSDGSNLEMLRTGCG 165
 QY 483 SPCYGRLEVYQGEWGTVDHWRSTRNAAVVCQKOLGCGKPMHVFMTYFKEASGPFIWDD 542
 DB 166 NMCSGRIEIKFGWRGTVDNDFNIDHASVICRQLECGSAVSFGSSNFGEGSGPIWFD 225
 QY 543 VSCIGNESINWCEHSGWCKNCHVREDVIVTCSDGATWGLRLVGSNRCSCRLVYFQG 602
 DB 226 LICNGNESALWCKHOGWKNCHDAEDAGVTCRSGADLSRLVDPVTECSGRLEVRFG 285
 QY 603 RWGTVCDDGWSKAAAVVCQKOLGCGPSSIIIGMGLGNASTGYGKIWLDDSCDGDDESILWSC 662
 DB 286 EWGTCDDGWSYDAVACKQLGCTPATAIGRVNASKGFGHIWLDVSCQCHPEAVWC 345
 QY 663 RNSGWNNDCHSEDVGVTCSDASDMELRVGSSRCAGKVEVNVQGVAGIICANGWGMN 722
 DB 346 KHEWKGKHCNHNEDAGVTCSDGSDLELRGGGRCAGTVEIOTRLGKVCDRGWLK 405

QY 723 IAEVVCVQLECGSAIRVSRPHFTERTLHILMSNSCGTGGASLWDCIRWEWKQTACHLN 782
 DB 406 EADVVCVQLECGSALTSYQYYSKIQTATNWLFLSSCNNGNETSLWDCNKMORGGTCDHY 465
 QY 783 MEASLICSARHROPRLVGADMPSCGRVEYKHADTWRSVCDSDFSLHAANVLNRELNCGDAL 842
 DB 466 EEAKITCSARHREPRLVGDDIPCSGRVEYKHGDTWGSICDSDFSLEAASVLCRELOGGTV 525
 QY 843 SLSVGDHFGKGNGLTAEKFCQEGSETHALCPVIOHPEDTCIHRSREYGVVCSRYTVDRL 902
 DB 526 SILGGAHFGEGNGQIWAEEFOCEGHEGSHSLCPVAPRPEGTCSHSDRVGVVCSRYTEIRL 585
 QY 903 VNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTGGKVIERSVRV 962
 DB 586 VNGKTPCEGRVELKTLGAWGLSCLNSHWDIEDAHVLCQQLKCGVALSTPGGARFGNGQI 645
 QY 963 WGRHFCILGNESLNDNCOMTVILGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 1022
 DB 646 WRHMEHCTGTQEHMGDCPVTALGASLCFSEQVAVSICSGNOSQTLSSCNSSSLGTRPTI 705
 QY 1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVCQKLCGGVA 1082
 DB 706 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGSGWTICDDSDWLSDAHVVCRLGCGGA 765
 QY 1083 ENATVSAHFEGSGPIWLLDNLCTGTEHLMWQPCPSRGWQHDCRHKEDAGVICSFEETALR 1142
 DB 766 INATVSAHFEGSGPIWLLDNLCTGTEHLMWQPCPSRGWQHDCRHKEDAGVICSFEFMSLR 825
 QY 1143 LYSETETESACGRLEVFYNGTWGVSRRNITTAIAGIVCRLQCGGNGVWSLAPLSKTGS 1202
 DB 826 LTSEASRACAGRLEVFYNGTWGVSRRNITTAIAGIVCRLQCGGNGVWSLAPLSKTGS 885
 QY 1203 GFMVDDIQCCKTHSIWQCSLAPWERRISSPAETWITCEDIRVIRVGDTSCSGRVEIW 1262
 DB 886 IPMWVDNVQCPKGPDTLWQCPSPWEKRLASPEETWITCDNKIRLQEGPTSCSGRVEIW 945
 QY 1263 HAGSGTVCDDSDWDLAEAVVCQKOLGCGSALAAALDASFGQGTGTIWLDDMKCNESFL 1322
 DB 946 HGSNGTVCDDSDWDLDAQVVCQKOLGCGCPALKAFAEFGQGTGTIWLNEVKCKNGESSL 1005
 QY 1323 WCHAKPMQSCDGHKEDAGVRCGQSLK-----SLASSGHLALILSSIFGLLLVLF 1376
 DB 1006 WDCPARRWGHSECGHKEADAANCTDISVQKTPQKATTGRSSRQSSFFIAVGLGVLLAIF 1065
 QY 1377 I--LFTWCRVOKHPLPLRVSTRRRGSLLENLFHEMETCLARE 1419
 DB 1066 VALFELTKRRQRQR---LAVSSRGENLVHQIYREMNCLNADD 1107

RESULT 8
 AAM41279
 ID AAM41279 standard; Protein; 1124 AA.
 XX
 AC AAM41279;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 6210.
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX Homo sapiens.
 OS
 XX WO200153312-A1.
 PW
 XX 26-JUL-2001.
 PD
 XX 26-DEC-2000; 2000WO-US34263.
 PF

25-APR-2000; 2000US-0552317.
09-JUL-2000; 2000US-0598042.
19-JUL-2000; 2000US-0620312.
03-AUG-2000; 2000US-0653450.
14-SEP-2000; 2000US-0682191.
19-OCT-2000; 2000US-0693036.
29-NOV-2000; 2000US-0727344.
(HYSE-) HYSEQ INC.
PA
XX
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI60436.
XX
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX
PS Example 2; SEQ ID NO 6211; 10078pp; English.
XX
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM3642-AAW4213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX
SQ Sequence 1124 AA;
Query Match 42.1%; Score 3429; DB 22; Length 1124;
Best Local Similarity 56.2%; Pred No. 2, le-238;
Matches 598; Conservative 161; Mismatches 294; Indels 12; Gaps 4;
364 GADLELRLADGNNCGSRVEVRIHEOWTTCIDQNMKNQALVYVCKQCPFFVFGSRRAK 423
Db 49 GTDKELRLVDGENKCSGRVEVVKQEBWGTVCNNGWSMEAVSVCNLQCGCPTAIAKAPGWAN 108
QY 424 PSNEARDIWNISCTGNESALWDCTYDCAKRT-CFRSDAGVTCSDKADLDLRLVGAH 482
Db 109 SSAGSGRIWHDVHSCRGNSALWDCKHDGKHSNCTHOODAGVTCSDGNSLEMLTRGG 168
QY 483 SPYGRLEVKYGEWTCVCHDRWSTRNAVAVCKQCGKPMHVFVGTKEASGPIWLD 542
Db 169 NMGSGRIEIKFQGWTKVTCDDNFNDHVASVTCQECGSAVSFGSSNFGSGPIWFD 228
QY 543 VSCIGNESNIWDCHEHSGWKHNCVHREDVIVTCSDATWGLRLVGSNRCGRLEVYFQG 602
Db 229 LICNGNESALWNCKHOGWKHNCDAEDAGVTCCKGADLSRLVDGVTCSGRLEVRFG 288
QY 603 RWTCTVDDGWSKAANVCSOLDPSSIIIGMGLGNASTCYGKTLWLDVSDCGDESPLWSC 662
Db 289 EWGTCDDGWDSDYDAVACKQJGCPFAVTAIGRVNASKGFGHIWLDVSCQGHAPWOC 348
QY 663 RNSGWNNDCHSEDEYGVICSDASDMELRLVGSRRSACGKVEYNVQAGVILCANGWN 722
Db 349 KHHEGWKHYCNHNDAGVTCSDGSDLELRLGGGSRACGTVEVEIQRIQLLKVCDRGWGLK 408
QY 723 IAEVCRQLECGSAIRVREPHETFTLHILMSNSCTGGEASLWDCIRWENKQFACHLN 782
Db 409 EADVWCRLQCGSALKTSQVYSKIOATNTWLPFLSSCNGNETSLWDCKNWQMGGLTCDHY 468

QY 783 MEASLICSARHQPRLVCGADMPGCRVEVKKHADTWRSVCDSDSFLHAANVLRELNCDAI 842
Db 469 EEAKITCSAHRPRLVGDIPCSGRVEVKGDTWGSICDSDSFLAAASVLRELQCCQFV 528
QY 843 SLSVGDHFGKNGLTWAEKFOCEGSETHALALCPVQIHPEDTCIHSREVGVVCSRYTDVRL 902
Db 529 SILGAHFGEGNGQIWAEEFOCEGSEHSHSLCLCPVAPREPCTCSHSRDVGVVCSRYTEIRL 588
QY 903 VNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCROLSCGTALSTTGGKYIGERSVRV 962
Db 589 VNGKTPCEGRVELATLGAWSGLCNHSHWDIEDAHVLCQQLKCGVALSTPGGARFKGNGQI 648
QY 963 WGRPHCLGNESLNDNCOMTVLGAPPCHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 1022
Db 649 WRHMFHCTGTQEHMGDCPVTALGASLCPSEQVAVSICSGNSQSLSSCNSSGLPTRTPT 708
QY 1023 PEGSALICLEDKRLRLVDGSRACGRVEIYHDGFWGTICDDGWDLSDAHVCQKLGCGVA 1082
Db 709 PEESAVACIESGQLKLVNGGRCGRVRIYHEGSGWTICDDSWDLSDAHVCRLQCGGEA 768
QY 1083 FNATVSAHFEGSGPIWLDLNLCTGTESHLWQCPSPRGWQHDCRHKEDAGVICSEFTALR 1142
Db 769 INATGSAHFEGSGPIWLDENKCNKESRIWQCHSHGWGQCNCRHKEDAGVICSFMSLR 828
QY 1143 LYSETETESACGRLEVFYNGTWGSRGNITATAGIYCROLGCCENGVS LAPLSKTGS 1202
Db 829 LTSEASREACAGRLEVFYNGTWGSRGNITATAGIYCROLGCCENGVS LAPLSKTGS 888
QY 1203 GFMWDDITQCPRKTHISIWQCLSAPEWERRISSPAAETWITCEDRIRVRGDTSCGRVEIW 1262
Db 889 IPMWVDNVQCPKGPDTLWQCPSPWEKRLASPEETWITCDNKIRLQEGPTSCGRVEIW 948
QY 1263 HAGSGWTVCDDSDWLAEBVVCQQLGCGSALAALRDASFQGTGTIWLDDMRCKNESFL 1322
Db 949 HGGSGWTVCDDSDWLDQAQVVCQQLGCGPALKAFAEFQGTGTIWLNEVKCKNESL 1008
QY 1323 WDCHAKPWGSDCGHKEDAGVRCGOSLK-----SLNASSGHLALILSSIFGLLLVLF 1376
Db 1009 WDCPARRMHSGHCEGHKEDAAVNCTDISVQKTPQKATGSSRSQSSFIAGVILGVVLLAIF 1068
QY 1377 I--LFTWCRVOKOKHPLRVSTRRRGSLLENLFHEMETCLKRED 1419
Db 1069 VALFELTKRRQRQ---LAVSSRGENLVHQIQYREMNNSCLNADD 1110
RESULT 10
AAU97584
ID AAU97584 standard; Protein; 1156 AA.
XX
XX AAU97584;
XX
XX 27-AUG-2002 (first entry)
XX
XX Human CD163 receptor cytoplasmic variant 2 protein sequence.
XX
XX Human: CD163; receptor; haptoglobin-haemoglobin complex; Hp-Hb complex;
KW CD163 receptor variant; antibiotic; anticancer drug; anti-HIV drug;
KW human immunodeficiency virus; haemolysis; haematological condition;
KW aplastic anaemia; iron deficiency anaemia; megaloblastic anaemia;
KW sickle-cell anaemia; polycythaemia; malaria; leukaemia; myelodysplasia;
KW lymphoma; leucopenia; splenectomy; inflammation; infection; cancer;
KW autoimmunity; immunodeficiency; CD163 receptor cytoplasmic variant 2.
XX
XX Homo sapiens.
XX
XX WO200232941-A2.
PN
XX
XX 25-APR-2002.
PD
XX
XX 12-OCT-2001; 2001WO-DK00671.
PF
XX
XX 16-OCT-2000; 2000DK-0001543.
PR
XX 11-JAN-2001; 2001DK-0000039.
PR

Domain 154..253
/label= Domain_2
/note= "Specifically claimed in claim 6"
Domain 261..360
/label= Domain_3
/note= "Specifically claimed in claim 6"
Domain 368..467
/label= Domain_4
/note= "Specifically claimed in claim 6"
Domain 473..572
/label= Domain_5
/note= "Specifically claimed in claim 6"
Domain 578..677
/label= Domain_6
/note= "Specifically claimed in claim 6"
Domain 714..814
/label= Domain_7
/note= "Specifically claimed in claim 6"
Domain 819..920
/label= Domain_8
/note= "Specifically claimed in claim 6"
Domain 924..1023
/label= Domain_9
/note= "Specifically claimed in claim 6"

WO200232941-A2.

25-APR-2002.

12-OCT-2001; 2001WO-DK00671.

16-OCT-2000; 2000DK-0001543.

11-JAN-2001; 2001DK-0000039.

22-FEB-2001; 2001US-270120P.

(PROT-) PROTEOPHARMA APS.

Moestrup S, Moller HJ;

WPI; 2002-452380/48.

New haptoglobin-haemoglobin complex or its mimics or receptors, useful in gene therapy, particularly for treating an individual suffering from e.g. haemolysis, anaemia, inflammation, infection, cancer or autoimmunity

Claim 1; Fig 5; 70pp; English.

The present invention relates to a new haptoglobin-haemoglobin (Hp-Hb) complex, or its part or mimic, being operably linked to a substance, capable of binding a CD163 receptor and/or a CD163 receptor variant. The Hp-Hb complex or its mimic, the CD163 variant or the CD163 receptor are useful as a medicament or in the manufacture of a medicament for treating an individual. The medicament comprises at least one drug to be delivered to a cell expressing a CD163 receptor or a CD163 variant. The medicament may also comprise at least one gene to be delivered to a cell expressing a CD163 receptor or a CD163 variant. The medicament is an antibiotic or an anticancer drug, e.g. an anti-HIV (human immunodeficiency virus) drug. In particular, the Hp-Hb complex or its mimic, the CD163 variant or the CD163 receptor, or the medicament are useful for treating haemolysis and/or other haematological conditions, aplastic anaemia, iron-deficiency anaemia, megaloblastic anaemia, sickle-cell anaemia, polycythaemia, malaria, leukaemia, myelodysplasia, lymphoma, leucopenia, splenectomy, inflammation, infection, cancer, autoimmunity or immunodeficiency. The method involves drug-delivery treatment or gene delivery treatment of an individual. The medicament is also useful for inhibiting (in vivo) uptake of haemoglobin in cells presenting a CD163 receptor. The CD163 variant is also useful in diagnosing, monitoring and/or controlling a condition in an individual. In particular, the CD163 variant is useful for identifying macrophages in a biological sample from an individual, for identifying at least one Hp-Hb complex in serum and/or plasma of an individual, or for removing at least one Hp-Hb complex in serum and/or plasma of an individual. The

CC present amino acid sequence represents the human CD163 receptor of the invention.

XX
SQ Sequence 1116 AA;

Query Match 42.1%; Score 3424; DB 23; Length 1116;
Best Local Similarity 56.1%; Pred. No. 4.8e-238;
Matches 597; Conservative 161; Mismatches 295; Indels 12; Gaps 4;

Qy 364 GADLELRADSGNCSGRVEVRIHEQWWTICDONWKNQALVVCVKQLGCPSPVFSRRRAK 423
Dy 41 GTDKELRLVDGENKCSGRVEVKVQEWCTVCGNNGMSMAVSVICNOLGCPPTAKPAGWAN 100
Qy 424 PSNEARDIWNISICTGNESALWDCYDYGAKRT-CFRRSDAGYICSDKADLDRLVGAH 482
Dy 101 SSAGSGRIWMDHVSCRGNESALWDCYDYGAKRT-CFRRSDAGYICSDKADLDRLVGAH 160
Qy 483 SPCYGRLEVKYQGEWGTCHDRWSTNRNAAVVCVKQLGCPSPVFSRRRAK 542
Dy 161 NMSGRIEIKFQGRWGTCDNFNIDHASVICRQLGCGSAVSFSGSSNFGEGSGPIWEDD 220
Qy 543 VSCIGNESNIWDCSHSGWGHKNCVHREDVIVTCSDATWGLRLVGGSNRCGRLEVFQG 602
Dy 221 LICGNESALWNCYQGEWGTCHDRWSTNRNAAVVCVKQLGCPSPVFSRRRAK 280
Qy 603 RWGTVCCDGMNSKAAAVVCVKQLGCPSPVFSRRRAK 662
Dy 281 EWGTVCCDGMNSKAAAVVCVKQLGCPSPVFSRRRAK 340
Qy 663 RNSGWNDCSHSGWGHKNCVHREDVIVTCSDATWGLRLVGGSNRCGRLEVFQG 722
Dy 341 KHEWGHKNCVHREDVIVTCSDATWGLRLVGGSNRCGRLEVFQG 400
Qy 723 IAEVVCRLGCGSALKTQVYYSKIQATNWLFLSSCNGNETSLWCKNMGWGLTCDHY 782
Dy 401 EADVVCRLGCGSALKTQVYYSKIQATNWLFLSSCNGNETSLWCKNMGWGLTCDHY 460
Qy 783 MEASLICSAPHRQPLVGDMPGSGRVEVKHADTWRSVCDSDSFLSAAVYLCRELNGDAI 842
Dy 461 EBAKITCSAPHRQPLVGDMPGSGRVEVKHADTWRSVCDSDSFLSAAVYLCRELNGDAI 520
Qy 843 SLVSGDHGKGNLTWAKFCQEGSETHALCPVQHPEDTCIHSRVEGVVCSRYTQVRL 902
Dy 521 SLVSGDHGKGNLTWAKFCQEGSETHALCPVQHPEDTCIHSRVEGVVCSRYTQVRL 580
Qy 903 VNGKSCQDQVEINVLGHWSLCTDHPEDARVLCRLQSLCTALSTTGKYGIGERSRV 962
Dy 581 VNGKSCQDQVEINVLGHWSLCTDHPEDARVLCRLQSLCTALSTTGKYGIGERSRV 640
Qy 963 WGRHFRHCLGNESLLDNQMTVLGAPPCIHGNTVSVLCGSLTQPLPCLANVSDPYLSAV 1022
Dy 641 WGRHFRHCLGNESLLDNQMTVLGAPPCIHGNTVSVLCGSLTQPLPCLANVSDPYLSAV 700
Qy 1023 PGESALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVQCKLQCGVA 1082
Dy 701 PGESALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVQCKLQCGVA 760
Qy 1083 FNATVSAHFEGSGPIWDLNCTGTESHLMQCPGRGWQCHDKHEDAGVTCSETAIR 1142
Dy 761 FNATVSAHFEGSGPIWDLNCTGTESHLMQCPGRGWQCHDKHEDAGVTCSETAIR 820
Qy 1143 LYSETETESACRLEVFYNGTWGVSRRNITTAIAGIVCRLQCGENGVVSLAPLSKTS 1202
Dy 821 LYSETETESACRLEVFYNGTWGVSRRNITTAIAGIVCRLQCGENGVVSLAPLSKTS 880
Qy 1203 GPMVDDIQCPTHTISWCLSPWERRISSPAETITCEDIRIRVGGDTSCSGRVEIW 1262
Dy 881 GPMVDDIQCPTHTISWCLSPWERRISSPAETITCEDIRIRVGGDTSCSGRVEIW 940
Qy 1263 HAGSGWTVCCDSDWDLAEAVVCVKQLGCGSALALRADASTGQGTGTITLWDMCKNGESFL 1322
Dy 941 HAGSGWTVCCDSDWDLAEAVVCVKQLGCGSALALRADASTGQGTGTITLWDMCKNGESFL 1000

QY 1323 WDCHAKPWGSDCGHKEDAGVRCGSQLK-----SLNASSGHLALILSSIFGLLLVLF 1376
 Db 1001 WDCPARWGHSECHKEDAAVNCCTDISVQKTPQKATTGRSSROSSFTAVGILGVLLAIF 1060
 QY 1377 I--LFLTWCRVQKHLPLRVSTRRGSLLENLFPHMETCLKRED 1419
 Db 1061 VALFFLTKRRQRQ---LAVSSRGLENLVHQIQYREMSNCLNADD 1102

RESULT 12
 AAU97583
 ID AAU97583 standard; Protein; 1149 AA.
 XX
 AC AAU97583:
 CC
 DT 27-AUG-2002 (first entry)
 XX
 DE Human CD163 receptor cytoplasmic variant 1 protein sequence.
 XX
 KW Human; CD163; receptor; haptoglobin-haemoglobin complex; Hp-Hb complex;
 KW CD163 receptor variant; antibiotic; anticancer drug; anti-HIV drug;
 KW human immunodeficiency virus; haemolysis; haematological condition;
 KW aplastic anaemia; iron-deficiency anaemia; megaloblastic anaemia;
 KW sickle-cell anaemia; polycythaemia; malaria; leukaemia; myelodysplasia;
 KW lymphoma; leucopenia; splenectomy; inflammation; infection; cancer;
 KW autoimmunity; immunodeficiency; CD163 receptor cytoplasmic variant 1.
 XX
 OS Homo sapiens.
 XX
 PN W0200232941-A2.
 XX
 PD 25-APR-2002.
 XX
 PF 12-OCT-2001; 2001WO-DK00671.
 XX
 PR 16-OCT-2000; 2000DK-0001543.
 PR 11-JAN-2001; 2001DR-0000039.
 PR 22-FEB-2001; 2001US-270120P.
 XX
 PA (PROT-) PROTEOPHARMA APS.
 XX
 PI Moestrup S, Moller HJ;
 XX
 DR WPI; 2002-452380/48.
 XX
 PT New haptoglobin-haemoglobin complex or its mimics or receptors, useful
 PT in gene therapy, particularly for treating an individual suffering from
 PT e.g. haemolysis, anaemia, inflammation, infection, cancer or autoimmunity
 XX
 PS Claim 1; Fig 5; 70pp; English.
 XX
 CC The present invention relates to a new haptoglobin-haemoglobin (Hp-Hb)
 CC complex, or its part or mimic, being operably linked to a substance,
 CC capable of binding a CD163 receptor and/or a CD163 receptor variant.
 CC The Hp-Hb complex or its mimic, the CD163 variant or the CD163 receptor
 CC are useful as a medicament or in the manufacture of a medicament for
 CC treating an individual. The medicament comprises at least one drug to be
 CC delivered to a cell expressing a CD163 receptor or a CD163 variant. The
 CC medicament may also comprise at least one gene to be delivered to a cell
 CC expressing a CD163 receptor or a CD163 variant. The medicament is an
 CC antibiotic or an anticancer drug, e.g. an anti-HIV (human
 CC immunodeficiency virus) drug. In particular, the Hp-Hb complex or its
 CC mimic, the CD163 variant or the CD163 receptor, or the medicament are
 CC useful for treating haemolysis and/or other haematological conditions,
 CC aplastic anaemia, iron-deficiency anaemia, megaloblastic anaemia,
 CC sickle-cell anaemia, polycythaemia, malaria, leukaemia, myelodysplasia,
 CC lymphoma, leucopenia, splenectomy, inflammation, infection, cancer,
 CC autoimmunity or immunodeficiency. The method involves drug-delivery
 CC treatment or gene delivery treatment of an individual. The medicament is
 CC also useful for inhibiting (in vivo) uptake of haemoglobin in cells
 CC presenting a CD163 receptor. The CD163 variant is also useful in
 CC diagnosing, monitoring and/or controlling a condition in an individual.

CC In particular, the CD163 variant is useful for identifying macrophages in
 CC a biological sample from an individual, for identifying at least one
 CC Hp-Hb complex in serum and/or plasma of an individual, or for removing at
 CC least one Hp-Hb complex in serum and/or plasma of an individual. The
 CC present amino acid sequence represents the human CD163 receptor
 CC cytoplasmic variant 1 of the invention.
 XX
 SQ Sequence 1149 AA;

Query Match 41.7%; Score 3396.5; DB 23; Length 1149;
 Best Local Similarity 54.4%; Pred. No. 4.8e-236;
 Matches 597; Conservative 161; Mismatches 295; Indels 45; Gaps 5;

QY 364 GADLELRADGNNCSGRVEVRHIEQWMTICDQWKNKEQALVVKQCGPFSVFGSRRAK 423
 Db 41 GTDKELRLVDGENKCSGRVEVKVQEEGTVCNNGMSMEAVSVICNLGCPAIRAPGAN 100
 QY 424 PSNEARDIWNISICTGNESALWDCTYDGAART-CFRSRDAGVICSADKADLRVLGAH 482
 Db 101 SSAGSGRIWMDHVSRCGNESALWDCKHDGKGNKSNCTHQDQAGVTCSDGSLNMLRTGG 160
 QY 483 SPYGRLEVKYQGEWGTVDHWRSTRNAAVVCKOLGCGKPMHVGMTVFEKASGPILWDD 542
 Db 161 NMSGRIEIKPQGRWGTVCDDNFIDHASVTCROLECGSAVSFGSSNFGGSGPIWFDD 220
 QY 543 VSCIENESNIWDCHSGWKGKNCVHREDVIVTCSDGATWGLRLVGGNRCGRLEVFQ 602
 Db 221 LICNGESALWNCQKQKGNKNCDAEDAGVICSAGDLSRLVDGVTECSGRLEVFQ 280
 QY 603 RWGTVCDDGWNKAAAVVCSQDPCSSIIGMGLNASTGYGKIWLDDVSCDDGSDLMSC 662
 Db 281 EMGTICDDGWDSDYDAAVACKOLGCPATAIGRVNASKGFGHILWDSVSCQGHPEAVWOC 340
 QY 563 FNSGWGNNDCSHSEDEVGVICSDASDMELRLVGGSSRCAGKVEYVVGAVGILCANGWGN 722
 Db 341 KHHEWKGHYCNHNEEDAGVTCSDGSDLELRGSGSACGTVEVEIQRLLGKVCYDRGWLK 400
 QY 723 IAEVVCROLECGSAIRVSRPHFTERTLHILMSNGCTGGGEASLWDCIRWEMKQATACHLN 782
 Db 401 EADVVCROLECGSALKTSYQVYSKIQTNTWLFSSCNGNETSLWDCNKNQWGGTLCDHY 460
 QY 783 MEASLICSAPRQLVGLADMPGSGRVEVKHADTWRSVCDSDSFLHAANVLCRELNCDAI 842
 Db 461 EEAKITCSAHPRLVGGDIPCSGRVEVKHGTWGSICDSDSFLHAANVLCRELNCDAI 520
 QY 843 SLSVCDHFGKNGLTWAEKFOCESETHLALCPVQHPEDTCIHSREGVVCSR----- 896
 Db 521 SILGAHFGEGNGQIWAEEFOCEGHESHLSCPVAPRPEGTCSHSDRVGVVCSKTKTKS 580
 QY 897 -----YTDVRLVNGKSCDQVQVEINVLGHWSICLDTHW 929
 Db 581 LIGSYTVKGTGLGSHSCLFLKPCLLPGYTEIRLVNGKTPCEGRVELKTLGAWGSLCNHW 640
 QY 930 DPEDARVLCRLSCGTALTSTGGKYIGRSVVRVNGHRPHCLGNBSLNDNCOMTVLGAPPC 989
 Db 641 DIEDAHLVCOOLKCGVALSTPGGARFGKNGQIWRHMFHCTGTQOHMGDCPVTALGSLC 700
 QY 990 IHGNTVSVICTGSLTLPCLANVSDPYLSAYPEGSALICLEDKRLRLVGDSCRACRV 1049
 Db 701 PSEQVAVSVCISGNSQTSLSNCSNSSLGPTPTPEESAVACIESGOLRLVNGGRCAGR 760
 QY 1050 EYHDGFWGTICDDGWDLSDAHVVCOKLGCQVAFNATVSAHFEGSGPIWLDLNLCTGE 1109
 Db 761 EYHEGSGWTICDDGWDLSDAHVVCRLGCGEATNATGSAHFEGGTGPIWLDENKCNKE 820
 QY 1110 SHLWQCPSRGQGHDRHKEDAGVICSSEFTALRLYSEFETESACARLEVFYNGTWGSYGR 1169
 Db 821 SRIWQCHSHGWQCNCRHKEDAGVICSSEFMSLRLTSEASREACARLEVFYNGAWGTGK 880
 QY 1170 RNITATAGVLCROLGCGGCVWSLAPLSKTCGFMWVDDIQCPKTHISIWQCLSAPOWER 1229
 Db 881 SSMSETTVGVVVCROLGADKGNIPASLDKAMSIPMWVYDQCPKPTDLWQCPSSPWEK 940

Db 886 SSMSETTVGVVCRQLGCKADKGINPASLDKAMSIPIWMVDNYQCPKGPDTLMQCPSPWEK 945

Qy 1230 RISSPAETWTCEDRIRVRGDTECSRVEIWHAGSNGTVCCDSDWDLAEAEVWCQQLGC 1289

Db 946 RLASPESETWTCDNKIRLQGGPTSCSRVEIWHGSGNGTVCCDSDWDLDDAQVVCQQLGC 1005

Qy 1290 GSALAALRDSFGQGTGTIWLDDMRCKNGESFLMDCHAKPMGQSDCGHKEBAGVRCSSQS 1349

Db 1006 GPALKAFKAEAFEGQGTPIWLNEVKCKNESSLMDCPARRNGSHSCGHKEBAAVNCNTDIS 1065

Qy 1350 LK-----SLNASSGHLALIUSSIFGLLLLVFI--LFTWCRVOKQKHLPLRVSTRRG 1401

Db 1066 VQKTPQKATTRSSQSSFIAGVILGVVLLAIFVALFELTKRRQROR---LAVSSRGEN 1122

Qy 1402. SLEENLFHEMETCLKRED 1419

Db 1123 LVHQIQYREMNLSCLNADD 1140

RESULT 14

AAB66088

ID AAB66088 standard; Protein: 1436 AA.

XX AC AAB66088;

XX DT 30-MAR-2001 (first entry)

XX DE Bovine WCI protein.

XX KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder; central nervous system; focal brain disorder; bipolar affective disorder; global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease; senile-dementia; Huntington's disease; amyotrophic lateral sclerosis; Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep; neuropsychiatric; psychoactive substance use; anxiety.

XX OS Bos sp.

XX PN WO200077239-A2.

XX PD 21-DEC-2000.

XX PF 24-MAY-2000; 2000WO-US14858.

XX PR 14-JUN-1999; 99US-0333159.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX DR WPI; 2001-032313/04.

XX DR N-PSDB; AAF45141.

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for screening assays and diagnostic assays and for the treatment of neurological diseases such as Alzheimer's, Parkinson's and Huntington's disease -

XX PS Disclosure; Fig 2; 359pp; English.

CC The present invention relates to TANGO or INTERCEPT proteins and coding sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057, AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding sequences are useful for the treatment of neurological disorders such as central nervous system (CNS) disorders, CNS-related disorders, focal brain disorders, global-diffuse cerebral disorders and other neurological and cerebrovascular disorders. The CNS disorders include Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome, autonomic function disorders such as hypertension and sleep disorders, neuropsychiatric disorders, psychoactive substance use disorders, anxiety, and bipolar affective disorder. The present sequence is a sequence used in a sequence homology alignment with the TANGO/INTERCEPT

CC sequences of the present invention.

XX SQ Sequence 1436 AA;

Query Match 37.9%; Score 3082; DB 22; Length 1436;

Best Local Similarity 41.7%; Pred. No. 3.le-213;

Matches 613; Conservative 208; Mismatches 495; Indels 154; Gaps 23;

Qy 28 CILLNSCFLLSSPNFTDLRLVNGDPCGCTVEVKFQGGMGTVCCDGNNTASTVVCCK 87

Db 13 CVLLIG-----TWGQQAELRLKDGVRHRCERVEKHQGGMGTVGRTWTLKUDASVVC 67

Qy 88 QLQCPFSFAM---FFRQAVTRHGKIWLDDVSCYNESALMECQH---REMGSHNYHGE 141

Db 68 QLQCGAAGFPGGAYFGPGL---GPIWLLYTSCECTESTVSDCEHSNIDKYENDCYNHGR 124

Qy 142 DGVNVCYGEANLGLRLVDGNNSCSRVEVKFOERNGTTCDDGNLNTAAVYCRQLGCPSS 201

Db 125 DAGVVCSG---FVRLAGGDGPCSGRVEVHSGEAMIPVSDGNFTLATQAIIICAEELGCGKA 180

Qy 202 FISSGVVNSPAVLRPIWLDDIILCOGNELALNCRHGRGNHDCSHNEDVTLCYDSSDLE 261

Db 181 VSVLGHLEFPRESSAQVWAEFRCEGEPELVPCPRVPGGTCCHHSGSAQVVCSEVR 240

Qy 262 LRLVGTNRNRCMGRVELKIQGRMGTVCHHKWNNAAADVVCQKQLGCGTALHFAGLPHLQSGS 321

Db 241 L-MTNGSSQCEQGVEMNISGQWRALCASHSANANVICRLGCGVAISTPGPHLVEEG 299

Qy 322 DVWLDGVSCSGNESFLMDCHSHGTVNFDLHONDVSVICSDGADLELRLADGNSNCSGR 381

Db 300 DOILTARFHCESAFLMSCPVTALGDPDCSHGNTASVICSGN----- 342

Qy 382 YEVRIHEQWTTICDQNKNEQALVVCQKQPCFPFVSGRRRAKPSNEARDIWNISICTGN 441

Db 343 -QIQVLPO-----CNDSV-----SQPTSAAS----- 363

Qy 442 ESALWDCITYDGKAKRTPCFRRSDAGVICSADKDLRLVGAHSPCYGRLEVKYQGGMGTV 501

Db 364 -----EDSAPYCSDSR--QLRLVDGGGPGCAGRVEILDQGSWGTIC 401

Qy 502 HDRWSTRNAAVVVCQKQCGKPMHVFMTYFKEASGPILWLDVSCIGNESNIWDCHSGWG 561

Db 402 DDGNDLDDARVVCRLGCGEALNATGSAHFGAGSGPILWLDNINCTGKESHVWRCPSRGW 461

Qy 562 KHCNVHREDVIVTCSDATWGLRLVGGSNRCSGRLEVFQGRWGTVCDDGWNSKAAAVVC 621

Db 462 QHNCRHQDAGVICS--EFLALRWVSEDOQCAGLEVYNGTWGVSVCNPNMEDIVSTIC 519

Qy 622 SOLDCPSSIIGMLGNAST---GYGKIWLDDVSCDGEDSLWSCNRSNGWNNDCSHSED 677

Db 520 ROLGCGDS---GTLNSSVALREGFRPWVDRIQCRKTDTSWQCPSPDPWYNNSCPKEE 575

Qy 678 VGVICSDASDMLRLVGGSSRCAGVENVQCAVILCANGHGMNIAEVVVCRLGCGSAI 737

Db 576 AYIWCADS--RQIRLVDDGGRCSGRVEILDQGSWGTICDDRWDLDDARVVCRLGCGEAL 633

Qy 738 RVSREPHTERTLHILMSNSGCTGCEASLWDCIRWEWQTCACHLNNEASLICASAHQRPL 797

Db 634 DATVSSFFGTGSGPTLWLDVENCREEEQVWRCPSPGWGRHNCNHQEDAGVCSGF--VRL 691

Qy 798 VGADMPCSRVEVKHADTWRSVCDSDFLSHAANVLCRELNCGDAISLSVGDHFGKNGUT 857

Db 692 AGGDGPCSGRVEVHSGEAWTPVSDGNFTLPTAQVICAELGCGKAVSVLGHMPFRESQGV 751

Qy 858 WAEKFCQCEGSETHLALCPVQHPEDTCHSRREVGVVCSRYTDVRLV--NGKSCDQGOVELN 916

Db 752 WAEERFCDDGEPPELWSCPVRVPCPGTCLHSGAAVVCVSYTEVQLMKNGTSQCEQVENK 811

Qy 917 VLGHWSGLCDTHWDPEDARVLCRLQSCGTALSTTGKVIIGERSVRVWGRHRCPLCNESLL 976

Db 812 TSGRRALCASHSANANVVCRLGCGVAISTPGPHLVEGGDQIQTAFQHCESAESFL 871

Qy 977 ONCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRL 1036

Db 872 WSPVYALGGPCSHGNTASVICSNGHTQVLPQCNDFLSQPAGSAASESSPYCSDSKQL 931
 QY 1037 RLVDGSRGAGVEIYHDFGWTICDDGDWDLSDAHVVCOKLGGVAFNATVSAHFGEGSG 1096
 Db 932 RLVDGGGCGGRVEILDQSGWTICDDGDWDLSDAHVVCOKLGGVAFNATVSAHFGAGSG 991
 QY 1097 PIWLDLNCITGTFESHLMQCPGRGWQHDCHRHKEDAGVICSEFTALRLYSETTESCAGRL 1156
 Db 992 PIWLDLNCITGTFESHLMQCPGRGWQHDCHRHKEDAGVICSEFTALRLYSETTESCAGRL 1049
 QY 1157 EYFNGTWSGVRGRNTTTTAAIGVRCQLCGGNGVVS LAPLSKTS GGFVWDDIQCPKTH 1216
 Db 1050 EYFNGTWSGVRGRNTTTTAAIGVRCQLCGGNGVVS LAPLSKTS GGFVWDDIQCPKTH 1109
 QY 1217 ISTWQCLSPWERRISPAETWITCE-----DRIRVRGGTECSGRVE 1260
 Db 1110 TSLWQCPSPGWKYSSCPKEAYISCEGRRPKSCPTAAACTDREKLRLGGSECSGRVE 1169
 QY 1261 IWHAGSWGTVCDSDWDLAEAEVVCQOLGCGSALALRDASFQGGTGTIWLDDMRCKGNBS 1320
 Db 1170 VWHGSMGTVCDDSWSLAEAEVVCQOLGCGSALALRDASFQGGTGTIWLDDMRCKGNBS 1229
 QY 1321 FLWDCHAKPWQSDCHGKEDAGVRCSG-----QSLKSLNASGHLALIL 1364
 Db 1230 SLWDCAEPWQSDCHGKEDAGVRCSGVRLTPTTTAGTTRTTSNLPGLFSLPVLCIL 1289
 QY 1365 SSIFGLLLVFLFTLCVQOKHLPLRVSTRRRRSLEENLFHEMETCLKREDPHGTR 1424
 Db 1290 GSLFLVLVLTQLRW-RAER-----RALSSYEDALAEAYEBLDYLLTQKEGLGSP 1342
 QY 1425 TS-DTTPNHCEDASDTSLGLVLPASEATK 1453
 Db 1343 DQMTDVPDENYDDAEVVPVPGTPSPSQGNE 1372

RESULT 15

AA64591
 ID AA64591 standard; Protein; 1785 AA.

AC AA64591;

XX 23-OCT-1998 (first entry)

DT Human SRCR protein.

DE Scavenger receptor cysteine rich domain; SRCR; diagnosis; treatment;
 KW nervous system; medullo-blastoma; glioma; breast; detection;
 KW antibody; ss.

OS Homo sapiens.

XX WO9830687-A2.

PN 16-JUL-1998.

XX 09-JAN-1998; 98WO-DE00096.

PF 18-JUL-1997; 97DE-1030997.

PR 09-JAN-1997; 97DE-1000519.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

PA Mollenhauer J, Poustka A;

XX WPI: 1998-399136/34.

DR N-PSDB; AA649652.

XX Proteins containing scavenger receptor, cysteine rich domain -
 PT useful for diagnosis and treatment of tumours

PS Claim 2; Fig 2; 54pp; German.

XX

CC This sequence represents a human protein which contains a SRCR (scavenger
 CC receptor, cysteine-rich) domain. The gene and encoded protein can be used
 CC to diagnose or treat tumours, particularly of the nervous system
 CC (medullo-blastoma or glioma) or breast. The DNA sequence and probes
 CC derived from it, are used to identify genes that express SRCR-domain
 CC containing proteins, to determine the form in which these proteins exist
 CC and to assess the significance of individual forms on cellular
 CC properties. The protein can be used to detect the presence of
 CC autoantibodies and antibodies which regulate its expression.

XX Sequence 1785 AA;

SQ Query Match 29.9%; Score 2435; DB 19; Length 1785;
 Best Local Similarity 36.1%; Pred. No. 1.9e-166;
 Matches 533; Conservative 155; Mismatches 415; Indels 374; Gaps 28;

QY 43 GTD--LELRVNGDPCSGTVEVKFGOWGTVCDDGNTTASTVVCVKQLGCPFSFAM--- 97
 Db 95 GDSGLALRLVNGDRCQGRVEILYRGSGTVCDDSWDTNDANVVCRLGCGWAMSAPON 154
 QY 98 FRFGQAVTRHGKIWLDDVSCYGNESALMPCQHRWGSHNCHYHGEDVYNCY----- 148
 Db 155 AWFQGG---SGPIALDDVRCSGHESYLWSCPHNGWLSHNCGHGEDAGVICSAQAPOSTLR 211
 QY 149 -----GEANLGLRLVDGNNCSGRVEVKEFQERWGTICDDGWNLTAAVY 192
 Db 212 PRSVPVRISSPPVTEGSESSALRLVNGDRCGRVEILYRGSGTVCDDYWDTDNDANV 271
 QY 193 CRLGCPSSFISSGVNNSPAVLRPIWLDDILCOGNELALWNCRHGNGHDCSHNEDVTL 252
 Db 272 CRLGCGWAMSAPGNAQFGQSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHSEDAGV 331
 QY 253 TCY-----DSSDLRLVGGTNRCMGRVELKTOGRWGTVCCHK 290
 Db 332 ICSPAQRPTSPDTPWTPTSHASTAGPESSALRLVNGDRCGRVEILYRGSGTVCDDSD 391
 QY 291 WNNAAADVVCVKQLGCGTALHFAGLPHLQSGSDVWVLDGVSCSGNESFLWDCRHSGTVNED 350
 Db 392 WDTSDANVVCRLGCGWATSAFNGARFGQSGPIVLDDVRCSGHESYLWSCPHNGWLSHN 451
 QY 351 CLHONDVSVICSDGADLELRADGNNCSGRVEVTHEQWWTICDQWNEQALVVCVKQL 410
 Db 452 CQHSEDAGVICSA-----HSWST----- 470
 QY 411 GCFPSVFGSRRAKPSNEARDIWNISCTGNESALWDCITYDGKAKRTCFRRSDAGVICSD 470
 Db 471 -----PSPDTPITLTPASTVGSSE----- 491
 QY 471 KADLDRLVGAHSPCYGRLEVKYQGEWGTVCCHDRWTRNAAVVCVKQLGCGKPMHVFGMTY 530
 Db 492 ---LALRLVNGDRCQGRVEILYQSGWTVCDSDWTDNDANVVCRLGCGWAMSAPGNAR 548
 QY 531 FKEASGPIWLDDVSCICIGNESNIWDCSHGKGNCHNCHREDVIVTCSGDA-----TW- 581
 Db 549 FGQSGPIVLDDVRCSGHESYPWSCPHNGWLSHNCGHSEDAGVICSAQSRPTSPDTPW 608
 QY 582 -----GLRLVGSNRCSGRLVEYFQGRWGTVCDDGWNNSKAAAVVCSQLDCPS 628
 Db 609 TSHASTAGSESSALRLVNGDRCQGRVEILYRGSGTVCDDYWDTDNDANVVCRLGCGW 668
 QY 629 SIIGMGLGNASTCYGKIWLDDVSCDGEDSLWSCRNSGNNDCHSDSEVDGVICS----- 683
 Db 669 AMSAPGNARFGQSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHEDAGVICSAQSQ 728
 QY 684 -----DASDMELRLVGGSSRCAGKVEYNVOGAVGILCANGMGMIAEY 726
 Db 729 PTPSPDTPWTSHASTAGSESSALRLVNGDRCQGRVEILYRGSGTVCDDYWDTDNDANV 788
 QY 727 VCRQLCEGSAIRVSRPHPTERTLHILMSNGCTGGEASLWDCIRWEWQKTACHLWNEAS 786
 Db 789 VCRQLGCGWATSAFNGARFGQSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHEDAG 848
 QY 787 LICSAHR-QPRLVGADMPSCGRVEVKHADTWRSVCDSDFSLSHAANVLCRELNCGDAISLS 845

```
Db      849 VICSASQSP-----TPSPDTWPT----- 867
Qy      846 VGDHFGKNGLTWAEKFCQEGSETHLALCPVQHPEDTCIHSREVGVVCSRYTDOVRLVNG 905
Db      868 -----SRAGSESTLAL-----RLVNG 886
Qy      906 KSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRLQSCGTALSTGGKYIGERSVRVWGH 965
Db      887 GDRCKGRVEVLYQSGWGTVCDDYWDTDANVVCRLCGWAMSAPNAQFGGSGPIVLD 946
Qy      966 RFHCLGNESLLDNQMTVLGAPCIHGNTVSVICTSLTO--PLFPCLANVSDPYLSAVP 1023
Db      947 DVRCSGHESYLWSCPHNGWLSHNGCHHEDAGVCSAAQSQSTPRPDTWLTTLNLPALTVGS 1006
Qy      1024 EGSALICLEDKRLRLVDGDSRCAGRVEIYHDFWGTICDDGWDLSDAHVVCOKLGCGVAF 1083
Db      1007 ESSL-----ALRLVNGDRCRVEVLYRGSWGTVCDDSDWTDNDANVVCRLCGCWAM 1059
Qy      1084 NATVSAHFEGSGPIWLDDLNCTGTESHLMQCPGSRGWGQHDCHRKEDAGVICSEFTALRL 1143
Db      1060 SAPGNAREFGSGPIVLDDVRCSGNESYLMSCPHKWLTHNCGHHEDAGVICS---ATQI 1116
Qy      1144 YSET-----ETESCAGRLEVY-NGTWGS----- 1166
Db      1117 NSTTTDWMHPTTTTTARPSSNCGGFL--FYASGTFSSPYYPYNNAKCVWEIEVANSY 1174
Qy      1167 ---VGRRI-----TTATAGVC---ROLGCGENGVSLL---AP 1196
Db      1175 RINLGFNLKLEAHNCSFDYVEIFDGSLSNLSLLGKICNDTROIFTSSYNRMTHIFRSD 1234
Qy      1197 LSKTSGGFM-WVDDIOCPKTHISIWQCLSAFWERRISSPAEETWITCEDRIRVRGGDTEC 1255
Db      1235 ISFQNTGFLAWYN-----SFPSDATL-----RLVNLNASSYCLC 1267
Qy      1256 SGRVEIHWAGSWGTVCDSDWDLAEAVVCOQLGCGSALAALRDASFGQGTGTIWLDDMRC 1315
Db      1268 AGRVEIYHGWTGTVCDSDSWTIQAEAVVCOQLGCGRAVSALGNAYFGSGGPIITLDDVEC 1327
Qy      1316 KGNESFLWDCHAKPWGQSDCGHKEKEDAGVRCGSQLKS 1352
Db      1328 SGTESTLWQCRNRCWFHNCNHRDAGVICSGNHLST 1364
```

Search completed: May 12, 2003, 06:37:17
Job time : 63 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 06:29:56 ; Search time 33 Seconds
(without alignments)
4232.828 Million cell updates/sec

Title: US-09-759-130B-381

Perfect score: 8138

Sequence: 1 MMLPQNSWHIDFGRCCHQNH.....CEDASDTSLGLVPASEATK 1453

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: piri:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3441.5	42.3	1151	2 I38004	M130 antigen precu
2	3429	42.1	1156	2 I38005	M130 antigen precu
3	3396.5	41.7	1149	2 I38006	M130 antigen precu
4	3082	37.9	1436	2 A46496	antigen WCL1.1 prec
5	2990	36.7	2403	2 A59386	sanko - human
6	2439.5	30.0	2153	2 T14893	scavenger receptor
7	2246.5	27.6	2043	2 T18524	scavenger receptor
8	2156.5	26.5	1594	2 T30549	hensin - rabbit
9	1962.5	24.1	2083	2 T42721	CRP-ductin-alpha p
10	1076.5	13.2	600	2 S56744	mucin (clone pgm7 -
11	1048.5	12.9	1290	2 A57190	ebnerin precursor
12	866.5	10.6	1036	2 T17405	scavenger receptor
13	720.5	8.9	761	2 JC5759	brain-specific ser
14	665	8.2	918	2 JC4361	scavenger receptor
15	626.5	7.7	626	2 I49100	mcd6 precursor -
16	594.5	7.3	468	2 S26741	T-cell glycoprotein
17	520	6.4	532	2 A32751	speract receptor p
18	385	4.7	504	2 S56745	mucin (clone pgm31
19	314.5	3.9	577	2 A53202	cyclophilin C-asso
20	311.5	3.8	585	2 A47161	Mac-2-binding glyc
21	290.5	3.6	518	2 A55840	macrophage bacteri
22	290	3.6	458	2 B44407	macrophage scaveng
23	288	3.5	451	2 A38415	macrophage scaveng
24	279.5	3.4	453	2 S08276	macrophage scaveng
25	279.5	3.4	454	2 I46862	macrophage scaveng
26	273.5	3.4	454	2 A44407	macrophage scaveng
27	264	3.2	494	1 A29079	lymphocyte surface
28	258	3.2	5376	2 T42215	zonadhesin - mouse
29	254.5	3.1	495	1 A26396	T-cell surface gly

cell-fate determin
Notch-1 protein -
fibrillin-1 precu
MEGF6 protein - ra
alpha-2-macroglobu
notch protein homo
T-cell surface gly
notch homolog - se
Babblani ring 3 pr
Notch homolog prot
fibrillin 1 precu
notch protein - fr
fibrillin I - bovi
alpha-2-macroglobu
fibrillin-2 precu
T-cell surface gly.

ALIGNMENTS

RESULT 1

I38004

M130 antigen precursor, splice form 3 - human

N:Alternate names: M130 antigen cytosolic variant 1

C:Species: Homo sapiens (man)

C>Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 17-Mar-2000

C:Accession: I38004; S36078

R:Law, S.K.; Micklem, K.J.; Shaw, J.M.; Zhang, X.P.; Dong, Y.; Willis, A.C.; Mason, D.
Eur. J. Immunol. 23, 2320-2325, 1993

A:Title: A new macrophage differentiation antigen which is a member of the scavenger

A:Reference number: I38003; MUID:93380506; PMID:8370408

A:Accession: I38004

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1151 <LAW>

A:Cross-references: EMBL:Z22969; NID:g312143; PIDN:CAA80542.1; PID:g312144

C:Comment: For other splice forms, see PIR:I34005 and PIR:I34006.

C:Genetics:

A:Gene: GDB:M130; GDB:MM130

A:Cross-references: GDB:119369; OMIM:133710; GDB:9958989

A:Map position: 10pter-10qter

C:Note: the locus has not been clearly established and there is some confusion in the

C:Superfamily: scavenger receptor cysteine-rich domain homology

C:Keywords: alternative splicing; glycoprotein; surface antigen; transmembrane protei

F:1-40/Domain: signal sequence #status predicted <SIG>

F:41-1151/Product: M130 antigen, splice form 3 #status predicted <MAT>

F:41-1045/Domain: extracellular #status predicted <EXT>

F:43-147/Domain: scavenger receptor cysteine-rich domain homology <SRC1>

F:151-254/Domain: scavenger receptor cysteine-rich domain homology <SRC2>

F:258-361/Domain: scavenger receptor cysteine-rich domain homology <SRC3>

F:365-468/Domain: scavenger receptor cysteine-rich domain homology <SRC4>

F:470-573/Domain: scavenger receptor cysteine-rich domain homology <SRC5>

F:575-678/Domain: scavenger receptor cysteine-rich domain homology <SRC6>

F:711-814/Domain: scavenger receptor cysteine-rich domain homology <SRC7>

F:816-920/Domain: scavenger receptor cysteine-rich domain homology <SRC8>

F:921-1024/Domain: scavenger receptor cysteine-rich domain homology <SRC9>

F:1046-1066/Domain: transmembrane #status predicted <TRN>

F:1067-1151/Domain: intracellular #status predicted <INT>

F:100,118,226,315,440,680,689,762,996/Binding site: carbohydrate (Asn) (covalent) #st

Query Match 42.3% Score 3441.5; DB 2; Length 1151;

Best Local Similarity 55.2% Pred. No. 2.le-215;

Matches 608; Conservative 166; Mismatches 311; Indels 17; Gaps 6;

Oy 364 GADLELRADGNNCSGRVEVRIHEQWMTICDQNKNEQALVQKLGCPFSVFGSRRAK 423

Db 41 GTDKELRLVDGKNCGRVEVKVQEGWCTVCNNGWMEAVSICNLGCPCTAIRKPGWAN 100

Oy 424 PSNEARDIWNISICTGNSALWDCTVDGKAKRT-CFRSRDAGVTCSDKADLDRLVGAH 482

Db 101 SSAGSGRIWMDHVHSCRGNSALWDCKHDMGKHSNCTHQDQAGVTCSDGSLNEMRLTRGG 160

Qy	483	S P C Y G R L E V K Y Q G E W G T V C H D R N S T R N A A V C V K O L G C G K P M H V F G M T Y T K E A S G P I W L D D	542
Db	161	N M C S G R I E K T F O G R W G T V C D N F N I D H A S V I C R Q L E G S A V S P S G S N F N G E G S G T W I F D D	220
Qy	543	V S C I G N E S N I W D C E H S G K N C V H R E D V I V T C S G D A T W G L R L U V G G S N C S R G L E V Y F O G	602
Db	221	L I C N G N E S A L W N C K H O G W G K H M C H A E D A G V I C S K G A D L S R L V D G V T C S G R L E V R F O G	280
Qy	603	R W G T V C D D G N S K A A A V C S Q L D C P S T I G M G L N A S T Y G K I W L D D V S C D G D E S D L W S C	662
Db	281	E W G T I C D D G W D S Y D A A V A C K O L G C P T A V T A I G R V N A S K G F H I W L D S V S C O G H E P A V M Q C	340
Qy	663	R N S G W G N D C S H E D Y G V I C S D A S M E L R L V G G S R C A G V E N V O G A V G I L C A N G M N	722
Db	341	K H W E K H Y C N N H D A G V T C D G S D L E R L V G G G S R C A G V E I Q L G L K V C D R G W G L K	400
Qy	723	I A E V V C R Q L E G C S A I R V S R E P H T E R T L H I L M S N G C T G G E A S L W D C I R W E W K Q T A C H L N	782
Db	401	E A D V V C R Q L G G S A L K T S Y Q V Y S K I O A T N T W L F L S S C N G N E T S L W D C K N W M G L T C D H Y	460
Qy	783	M E A S L I C S A H R Q P L V G A D M P C S G R V E K H A D T W R S V C D S F S L H A N V L C R E L N G C D A I	842
Db	461	E E A K I T C S A H R P L V G G D I P C S G R V E K H G D T W G S I C D S F S L E A S V L C R E L Q C G T V V	520
Qy	843	S L S V G D H F G K N G L T W A E F Q C E G S E T H L A L C P I V Q H P E D T C I H S R E V G V V C S R Y T D V R L	902
Db	521	S I L G A H F G B G N G O I W A E F O C E G H E S H L S L C P V A P R P E G T C S H R D V G V V C S R Y T E I R L	580
Qy	903	V N G S Q C D G O V E I N V L G H G S L C D T H W P D E A V L C R O L S C G T A L S T T G T G K Y I G E R S V R	962
Db	581	V N G T P C E G R V E L K T L G A M G S L C N H W D I E D A H V L C Q Q K C V A L S T P G A R G K G N G Q I	640
Qy	963	W G H R F L C G N E S L D N C M T V L G A P P C T H G N T V S V I C T G S L T Q P L F C L A N V S D P Y L S A V	1022
Db	641	W R H M F H C T G T E Q H M G D C P V T A L G A S L C P S E Q V A S V I C S G N S Q T L S S C N S S L G P T R P T I	700
Qy	1023	P E G S A L T C L E D K R L R L V D G D S R C A G R V E Y H D G W G T I C D D G W D L S A H V V C O K I G C G V A	1082
Db	701	P E E S A V A C I E S G Q L R L V G G R C A G R V E Y I H E S W G T I C D D S W D L S A H V V C R Q L G C G E A	760
Qy	1083	F N A V S A H F G E G S P I W L D D L N C T G T E S H L W O C P S R G W G Q H D C R H K E D A G V I C E F T A L R	1142
Db	761	I N A T G S A H F G E G T P I W L D E M K N G E S R I W O C H S H G W G Q O N C R H K E D A G V I C E F M S L R	820
Qy	1143	L Y S E T E T S C A G R L E V Y N T W G S V G R R N I T T A J A I V C R Q L G C G E N V Y S L A P L S T G S	1202
Db	821	T L S E A S R E A C A G R L E V Y N A M G T V G K S M S E T T V G V V C R Q L G C A D K G K I N P A S L D K A M S	880
Qy	1203	G F M V D D I Q C P K T H I S T W O C L S A P E R R I S S P A E E T W I T C B E D R I K R V G D T E C S G R V E I W	1262
Db	881	I P M V D N V Q C P K G P D T L W C P S P S P E K R L A S P S E E T W I T C N K I R L Q E G P T S C S G R V E I W	940
Qy	1263	H A G S G T V C D D S W L A E A E V V C O O L G C G S A L A L R D A S F G Q C T G P I W L D M R K G N E S F L	1322
Db	941	H G S W G T V C D D S W D L D A Q V V C O O L G C G P A L K A F E A F E G Q G T G P I W L N E V K K C N E S S L	1000
Qy	1323	W D C H A K P W G S D C G H K E D A G V R G S G L K - - - - - S L N A S G H L A L I L S S I F G L L L L V L F	1376
Db	1001	W D C P A R R W G H S E C G H K E D A A V N C T D I S V O K T P O K A T T G R S S Q S S F I A V G I L G V L L A I F	1060
Qy	1377	I - - - - - L F L T W C R V O K H L P L A V S T R R G S L E E N L F H E M E T C L K R E D P H G T R T S D D T P N H C	1434
Db	1061	V A L F F L T K R K R O R - - - - - L A V S R G E N I V H I Q I Y R E M N S C L N A D D L D M N S E N - S H E S	1115
Qy	1435	E D A S D T S L G V - - - - - L P A S E A T K 1453	
Db	1116	A D F S A A F I I S V S K F L P I S G M E K 1137	

C:Species: Homo sapiens (man)
C:Date: 17-May-1996 sequence_revision 17-May-1996 #text_change 17-Mar-2000
C:Accession: I38005; S36079
R:Law, S.K.; Micklem, K.J.; Shaw, J.M.; Zhang, X.P.; Dong, Y.; Willis, A.C.; Mason, D.
Eur. J. Immunol. 23, 2320-2325, 1993
A:Title: A new macrophage differentiation antigen which is a member of the scavenger
A:Reference number: I38003; MUID:93380506; PMID:8370408
A:Accession: I38005
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1156 <LAW>
A:Cross-references: EMBL:222970; NID:g312145; PIDN:CAA80543.1; PID:g312146
C:Comment: For other splice forms, see_PIR:I34004 and PIR:I34006.
C:Genetics:
A:Gene: GDB:M130; GDB:MM130
A:Cross-references: GDB:119369; OMIM:133710; GDB:9958989
A:Map position: 10pter-10qter
A:Note: the locus has not been clearly established and there is some confusion in the
C:Superfamily: scavenger receptor cysteine-rich domain homology
C:Keywords: alternative splicing; glycoprotein; surface antigen; transmembrane protein
F:1-40/Domain: signal sequence #status predicted <SIG>
F:41-1156/Product: M130 antigen, splice form 4 #status predicted <MAT>
F:41-1156/Domain: extracellular #status predicted <EXT>
F:43-1045/Domain: scavenger receptor cysteine-rich domain homology <SRC1>
F:151-254/Domain: scavenger receptor cysteine-rich domain homology <SRC2>
F:258-361/Domain: scavenger receptor cysteine-rich domain homology <SRC3>
F:365-468/Domain: scavenger receptor cysteine-rich domain homology <SRC4>
F:470-573/Domain: scavenger receptor cysteine-rich domain homology <SRC5>
F:575-678/Domain: scavenger receptor cysteine-rich domain homology <SRC6>
F:711-814/Domain: scavenger receptor cysteine-rich domain homology <SRC7>
F:816-920/Domain: scavenger receptor cysteine-rich domain homology <SRC8>
F:921-1024/Domain: scavenger receptor cysteine-rich domain homology <SRC9>
F:1046-1066/Domain: transmembrane #status predicted <TRN>
F:1067-1156/Domain: intracellular #status predicted <INT>
F:100-119,232,335,440,580,689,762,996/Binding site: carbohydrate (Asn) (covalent) #st

Query Match	42.1%	Score 3429;	DB 2;	Length 1156;
Best Local Similarity	56.2%;	Pred. No. 1.3e-214;		
Matches	598;	Conservative 161;	Mismatches 294;	Indels 12; Gaps
QY	364	GADLELRADSNCSGRVEVRHEOWMTICDONWKEQALVYCKQLGCPFSVFGSRRAK	423	
DB	41	GYDKELRLVDGCKSGRVEVKVQEEWGTVCNNGWSMEAVSVCNLQGCPTALKAPGANW	100	
QY	424	PSNEARDIWINISICTGNESALMDCTYDGGAKRT-CFRSRDAGVTCSDKADLDRLVGNH	482	
DB	101	SSAGSGRIWMDHVSRCGNESALWDCKHGDKGHKSNCTHQODAGVTCDSGNLEMRTRGG	160	
QY	483	SPCYGRLEVKVQEWGTVCHDRWSTRNAAVYCKQLGCGRPMHVFGMTYFKEASGPIWLDD	542	
DB	161	NMCSGRIEIKFGRWGTVCDNFNIDHIASTICRLECGSVAVSFGSSNFGESGPIWFFD	220	
QY	543	VSCIGNESINWDCSHSGWGKNCVHRVEDVITVCSGATWGLRVLGSGNSCRLEVFYFG	602	
DB	221	LICNGESALMNCKHQGWKKINCDBAEDAGVICSKGADLSRLVDGVTECSGRLEVRFG	280	
QY	603	RWGTVCDDGWSKAAAVYCSQLDCPSSIIGMGLGNASTYGGKTIWLDDVSCDGDSDSLWSC	662	
DB	281	EWGTCDDGWSYDAAVACKLGCPTAVTAIGRVNASKFGFIWLDSVSCQGEHPAVWC	340	
QY	663	RNSGWNDCDSHSDGVVICSDASDMLRLVGGSSRCAGKVEVNVQAGVILLCANGWGN	722	
DB	341	KHEWGHKYCNHEDAGVTCSDGSDLELRLVGGSGRCAGTVEIQRLLKGVCDRWGLK	400	
QY	723	IAEVVCRQLCGSALRVSRPEHFETRLHLILMSNGCTGGEASLWDCIRWENKQTAACHLN	782	
DB	401	EADVVCRLQJGCSALKTSYQISKIQTATNWLFLSSCNGNETSLWDCCKNMQWGGLTCDHY	460	
QY	783	MEASLITCSAHRPRLVGADMPSCSRVEVKHADTWRSVCDSDFSLHAANYLRLCNGCDAI	842	
DB	461	EBAKTCASHREPRLVGGDI PCSGRVEVKHGDWTGSDSDFSLEAAVSLCRLCQGVTV	520	
QY	843	SLSVGDHFGKGNGLTWAEKFCQESGETHALCPIVQHPEDTCHSREVGVGVCSEYTDVRL	902	

[illegible]

RESULT 4

A46496
antigen WCL1.1 precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 18-Jun-1993 #sequence revision: 19-May-1994 #text_change 05-Nov-1999
C:Accession: A46496; I45834; S19913
R:Wijngaard, P.L.; Metzelaar, M.J.; Machugh, N.D.; Morrison, W.I.; Clevers, H.C.
J. Immunol. 149, 3273-3277, 1992
A:Title: Molecular characterization of the WCL antigen expressed specifically on bovine
A:Reference number: A46496; MUID:93056489; PMID:1431105
A:Accession: A46496
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1436 <W1>
A:Cross-references: EMBL:X63723; NID:g13; PIDN:CAA45255.1; PID:g14
A:Experimental source: CD4-CD8- gamma delta T lymphocytes
A:Note: sequence extracted from NCBI backbone (NCBIP:1117475)
R:Wijngaard, P.L.; Machugh, N.D.; Metzelaar, M.J.; Romberg, S.; Bensaid, A.; Pepin, L.;
J. Immunol. 152, 3476-3482, 1994
A:Title: Members of the novel WCL1 gene family are differentially expressed on subsets of
A:Reference number: I45834; MUID:94194107; PMID:7511649
A:Accession: I45834
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1436 <W1>
A:Cross-references: EMBL:X63723; NID:g13; PIDN:CAA45255.1; PID:g14

[illegible]

```
|||||
Db 1050 EVFYNGTWSVCRSPMEDITVSVICRQLGCGSLNTSVGLREGSRPRWVLDIOCRMD 1109
Qy 1217 ISIHQCLSAWERRISPAETHWTCB-----DRIVRGDDTECSGRVE 1260
Db 1110 TSLWQCPGPMKYSSCKPEAYISCERRPKSCPTAAACTDREKRLRGDSECSGRVE 1169
Qy 1261 IWHAGSNGTVCCDSDWDLAEAEVVCQQLGCGSALAALDASFGQGTGTIWLDDMRCKNGES 1320
Db 1170 VWHNGSNGTVCCDSDWDLAEAEVVCQQLGCGSALAALDASFGQGTGTIWLDDMRCKNGES 1229
Qy 1321 FLMDCHAKPMQSDCGHKEDAGVRCG-----QSLKSLNASSGHLALIL 1364
Db 1230 SLMDCAEPMQSDCKHEEDAGVRCGVRITLPTTAGTRTTSNLSGIFSLPGVLCLIL 1289
Qy 1365 SSIFGLLLVILFLFWCRVOKHPLRVSTRRRSLEENLHEMETCLKREDPHCTR 1424
Db 1290 GSLFLVILVLTQLRW-RAER-----RALSSYEDALAEAVYEELDYLLTQKEGLGSP 1342
Qy 1425 TS-DDTPNHGCEADSDTSLGLVLPASEATK 1453
Db 1343 DQMTDVPDENYDAAEVPVPGTSPSQONE 1372

RESULT 5
A59386
sanko - human
C:Species: Homo sapiens (man)
C:Date: 31-Dec-2001 #sequence_revision 31-Dec-2001 #text_change 31-Dec-2001
C:Accession: A59386
R:sanko, S.
submitted to the Protein Sequence Database, March 2001
A:Reference number: A59386
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-2403 <SAN>

Query Match 36.7%; Score 2990; DB 2: Length 2403;
Best Local Similarity 40.5%; Pred. No. 9.2e-186;
Matches 619; Conservative 177; Mismatches 504; Indels 230; Gaps 23;

Qy 44 TDLELRVNGDGPCTGVVEKFGQWGTVCDDGNNTASTVVKOLGCPFSFAM---FRF 100
Db 230 SSLALRLVNGDGRGRVLYRGSGTVCDDYWDNDANVVCRLGCGWAMSAPGNAQF 289
Qy 101 GOAVTRCKIWLDDVSCYGNESALWECQREHNGSHNCYHGEDVGVNCYG----- 149
Db 290 GQG---SGPIVLDVRCSGHESYLWSCPHNGWLNHNGHSEDAGVICSAPLSRPTSPDT 346
Qy 150 -----EAMGLRLVDGNNSCSGREVVEKFGQWGTVCDDGNNTASTVVKOLG 198
Db 347 WPTSHASTAGPESLALRLVNGDGRGRVLYRGSGTVCDDSDWTDNDANVVCRLG 406
Qy 199 PSSIFSGVNSPAVLRIWLDDILCOGNELALWNCRRHNGHDCSHNEDVTTCYDS- 257
Db 407 GWATSAPGNARFGQSGPIVLDVRCSGYESYLWSCPHNGWLNHNGHSEDAGVICS 466
Qy 258 -----SDLELRVNGTNRNCRGRVELAIOGRWGTVCCHKNNAADVVCQLG 304
Db 467 PTITLPASTVGSSESLALRLVNGDGRGRVLYRGSGTVCDDSDWTDNDANVVCRLG 526
Qy 305 CGTALHFAGLPHLQSGSDVWLDCVSCGNEFSLWDCRHSTGVNFDCLOHNDVSVICS-D 363
Db 527 CGWAMLAPGNARFGQSGPIVLDVRCSGHESYLWSCPHNGWLNHNGHSEDAGVICS 586
Qy 364 GADLELRDLADSNNSCRGVEIRIHEQWNTICDQNKNEQALVVKOLGCPFSVFGSRAK 423
Db 587 ESSLALGLVNGDGRGRVLYRGSGTVCDDSDWTDNDANVVCRLGCGWATSAPGNAR 646
Qy 424 PSNEARDIWNISICTGNESALWDCTYDGKAKRTCFRRSDAGVICS----- 469
Db 647 FGQSGGPIVLDVRCSGHESYLWSCPHNGWLNHNGHSEDAGVICSAAQSRSTRPDTLS 706
```

```
Qy 470 -----DKADLDRLVCAHSPCYGRLEVKYQEGMGTVCHDRWSTRNAAVVC-KOLGC 519
Db 707 TITLPPSTVGSSESLTLRLVNGSDRCQGRVEVLRGSHGTVCDDSDWTDNDANVVCRLG 766
Qy 520 GKPMHVFGMTYFKKASGPIWLDDVSCIGNESINWDEHSGKNGKNCVHREDVITCS--- 576
Db 767 GWATSAPGNARFGQSGPIVLDVRCSGHESYLWSCPHNGWLNHNGHSEDAGVICSVSQ 826
Qy 577 -----GDAW-----GLRLVGSNRCSGRLEVYFQGRWGTVCDDGHNKAA 617
Db 827 SRPTSPDPTWPTSHASTAGSESLALRLVNGDGRGRVLYRGSGTVCDDSDWTDSDA 886
Qy 618 AVCSQDLDCPSIIGMGLGNASTGVKILWDDVSCDGEDSLWSCRNSGNGNDCSHSED 677
Db 887 NVCCRLCGGWATAPGNAREFGQSGPIVLDVRCSGYESYLWSCPHNGWLNHNGHSED 946
Qy 678 VGVTCSDA-----SOMELRLVGGSSRCAGKVEVNVQAGVI 713
Db 947 AGVICSAAHSMSTPSPDTLPTITLPASTVGSSESLALRLVNGDGRGRVLYRGSGT 1006
Qy 714 LCANGMGNIAEVVCRLECGSALRVSRPHFTERTLHLMSNGCTGEASLMDIRWE 773
Db 1007 VCDSDWTDNDANVVCRLGCGWAMSAPGNARFGQSGPIVLDVRCSGHESYLWSCPHNG 1066
Qy 774 WKOTACHLNMEASLITCSAHR-QP-----RLVGADMPGCSGRVE 809
Db 1067 WLNHNGHSEDAGVICSASQSRPTSPDPTWPTSHASTAGSESLALRLVNGDGRGRVE 1126
Qy 810 VKHADTVRSVCDSDFSLHAANVLCRELNGDAISLVSDGHFGKGNGLFWAKFKCEGET 869
Db 1127 VLYRGSGTVCDDYWDNDANVVCRLGCGWAMSAPGNAREFGQSGPIVLDVRCSGHES 1186
Qy 870 HIALCPVQHPEDTCIHSREVGVCRYTD-----VRLVNG 905
Db 1187 YLWSCPHDGLWLNHNGHSEDAGVICSASQSRPTSPDPTWPTSHASTAGSESLALRLVNG 1246
Qy 906 KSCQDGOVEINVLGHWSLCTDTHWDPEDARVLCRLSCGTALSTTGGYIGERSVRVWGH 965
Db 1247 GDRQGRVLYRGSGTVCDDYWDNDANVVCRLGCGWATSAPGNARFGQSGPIVLD 1306
Qy 966 RFHCLGNESLNDNCOMTVLGAPPCHIGNTVSVICTGTSITQPLFPCLANVSDPYLSAVPEG 1025
Db 1307 DVRCSGHESYLWSCPHNGWLNHNGHSEDAGVICSASQSRPT-PSPTWPTSHASTAGSE 1365
Qy 1026 SALICLEDKRLRLVDGSRACAGRVEIYHDGFWGTICDDGWLSDAHVVCOKLGCQVAFNA 1085
Db 1366 SSL-----ALRLVNGDGRGRVLYRGSGTVCDDYWDNDANVVCRLGCGWATS 1419
Qy 1086 TVSAHFGEQSGPIWLDDILNCTGTETSHLWQCPSRGQGHDCRHKEDAGVICS----- 1135
Db 1420 PGSARFGQSGPIALDDVRCSGHESYLWSCPHNGWLNHNGHSEDAGVICSASQSRPTPS 1479
Qy 1136 -----SEFT-ALRLYSETETESACGRLEVEYNGTWGSGVRNRTTAIAGIVC 1181
Db 1480 PDTWPTSRASTAGSESLALRLVNGDGRGRVLYRGSGTVCDDYWDNDANVVC 1537
Qy 1182 ROLGCGENGVSAPLSK---TGSCFMVDDIQCPTHISTWOCLSAPW----- 1227
Db 1538 ROLGC---GWAMAPGNAQFGQSGPIVLDVRCSGHESYLWSCPHNGWLNHNGHSED 1594
Qy 1228 -----ERRISPAETHWTCB-----CEDRIRVR---GDDTECSGRVEIHWAGSGT 1269
Db 1595 GVICSAAQSGTSPDPTWLTNLPALTYGSESLALRLVNGD-RCRGRVLYRGSGT 1653
Qy 1270 VCDSDWDLAEAEVVCQQLGCGSALAALDASFGQGTGTIWLDDMRCKNGESFLMDCHAKP 1329
Db 1654 VCDSDWTDNDANVVCRLGCGWAMSAPGNAREFGQSGPIVLDVRCSGHESYLWSCPHNG 1713
Qy 1330 WQSDCGHKEADAGVRCSCQSLKSLNASSGH 1359
Db 1714 WLTHNCGHSEDAGVICSATQINSTTDDWWH 1743
```

RESULT 6

T14893
scavenger receptor cysteine-rich protein precursor - sea urchin (Strongylocentrotus purpuratus)
C:Species: Strongylocentrotus purpuratus (purple urchin)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
C:Accession: T14893
R:Pancer, Z.; Rast, J.P.; Davidson, E.H.
Immunogenetics 49, 773-786, 1999
A:Title: Origins of immunity: transcription factors and homologs of effector genes of the sea urchin
A:Reference number: 218253; MUID:99328904; PMID:10398804
A:Accession: T14893
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2153 <PAN>
A:Cross-references: EMBL:AF064259; NID:g4165052; PID:g4165053; PIDN:AAD08654.1
C:Genetics:
A:Gene: SRCR12

Query Match 30.0%; Score 2439.5; DB 2; Length 2153;
Best Local Similarity 38.2%; Pred. No. 4.3e-150;
Matches 527; Conservative 197; Mismatches 560; Indels 97; Gaps 22;
QY 32 LNSCFLLISFN-----GTDLELRVNGDPCSGTVEYKFGQGWCTVCCDGN 78
DB 113 LSSCAVSSHNCAHSEDAGVWCAGSGDGNIRLVGSHSEGRVELYVNNQWTVCCDLWD 172
QY 79 TTASTVVCCKLGC-PFSAMREFGOAVTRHGKIWLDDSCYCNESALWECQREWGSHNC 137
DB 173 LNDQVACRQLGLGPAVSADAFEGE---SGSILLDNVACITETISLLCSHPGIESHNC 229
QY 138 YHGEDVGNVCYGEANLGLRLVDGNNSCGRVEYKFGQGWCTVCCDGN 197
DB 230 GHSEDAGVTCSSRFSGDIRVVGSGIPTGREGVEYFVNGAMGTVCDDLWDINDASVACRQLG 289
QY 198 CPSSFTSSGVNVPVLRPIWLDDILCOGNELALNCRHNGHNDCHNEVDVLTCTYDS 257
DB 290 FGRTASGAGSYGGSGSIFLDNLACTGAESNLLSCPHNGVGSNCHGEDAGVLCSSS 349
QY 258 SDELRVGTGTRNCRBELKIQGRWGTVCCHKNAAADVVCCKQLGCGTALHFAFLPHL 317
DB 350 LSSDIRVVGGRSPTEGREGVEYFVNRANGTVCCDLWDVNDVNVACRQLGFGRAISAPGASY 409
QY 318 QSGSDVWLDDVSCSGNESFLMDCRHSHTVNFDCILHNDVSVICSDGA--DLELRADGNS 376
DB 410 GQSGSILLDNLACTGAESILLSCPHNGVGSNCHGEDAGVSCAPSSQESRVRVLGGLN 469
QY 377 NCSGRVEYRIHQWWTICDQNKNEQALVVCCKQLGCPFPFVFGSRRAPK--SNEARDIWLN 434
DB 470 NREGVEIFLNQWGTVCDDDDGTPDANVVCCKQLGTPSG--GSARSAVFGSGVPILLD 527
QY 435 SISCCTGNESALMDCTYDGRKAKRTCFRRSDAGVTC---SDKADLDRLVGAHSPCYGRLEV 491
DB 528 NVCCSGNERSLELCSNNGVHNCGHQEDASVVCVTGVTMPGMNIRIVGGSPSEGRREV 587
QY 492 KYOGEVGTCHDRWTRNAAVVCCKQLGCKPMHVFQMTYFKASGPILWDDVSCITGNESN 551
DB 588 LVGHRWGTCCDLWDINDANVVCRELGYSAATSATSSAFSGQSGGQILLDLRCSSTESS 647
QY 552 INDCBSHGKNCVHREDIVVTCSDATWG-----LRLVGSNRCGRLEVIYFQGRWG 605
DB 648 LITCPHGVGVHNCASEDAGVVCA--SITSGPVGVGYMRLVGGQNSRQGRLEISINNQG 706
QY 606 TVCCDGNWNSKAAVVCQLDPCSSITGMGLGNASTCYGKIWLDDVSCDGEDSLWNSCRNS 665
DB 707 TVCCDSDINDATVVCCKQLGFSAPVSAFTSAHFGQSGGTIWLDDVSCAGNENSLMDCGHR 766
QY 666 GWNNDCHSHEDVGVTCSDASD--MELRLVGSNRCAGKVEVNVQCAVILCANGWGMNI 723
DB 767 GLGVHNCACHAEDAGVYCI--ASDGLPIRLAGRSGMEGRVEISLGGDWGTVCCDSDGIED 825
QY 724 AEVVCCKQLGCGSAIRVSRPHFTERTLHILMSNGCTGCGEASLWDCIRWENKQTACHLNN 783

DB 826 AHVVCCKQLGFGPALSAYTAASFQGGSGSILMDNVQCSGDEATIAPCSHNGIGHNCGHQE 885
QY 784 EASLICS-AHRQPRLVGADMPSCSRVEYKHAHTWRSVCDSESLHAANVLRELNCDAI 842
DB 886 DAGVVCSSRATGEVRLVGNRSQNEGRVEILISGRWGTVCDDLWDLRADADVICQLGYGNAI 945
QY 843 SLSVGDHFGKGNLTWAEKFOCEGSETHALCPIVQHP---EDTCIHSREVGVCVSRVTD 899
DB 946 SAPHSSFFGPGRGDILLDDVSCGSEDDILDC---SHPIGTNNCGHSEDAGVCCD--LN 1000
QY 900 VRLVNGKSCQDQVQEVINVLHGWSLCTDHPEDARVLCRQLSCGTALTSTGGKYIGERS 959
DB 1001 VRIVNGSRNENGRVEILLHDGSHWATVCCDNDWLDATVVCCKQLGFPFSYAVAHAVFGGGS 1060
QY 960 VRYVNGHREHFCNLGNESLLDNCOMTVLGPAPCIHNGTVSVICTGSLTQPLPCLANVSDPYL 1019
DB 1061 GEIVLDDVECTGDEVSLETCQHAGLGTNNCGHSEDAGVICS-----VNV----- 1104
QY 1020 SAVPEGSSALICLEDKRLRLVDGDSRCAGRVEIYHDFWGTICDDGMDLSDAHVVCCKLGC 1079
DB 1105 -----RLADGNSPAEGRVEYFDGQWGTVCDDNMGWQNGHVICRAVGF 1147
QY 1080 GVAFNATYSAHFEGSGPIWLDDLACTGTESHLWOCPSRGWQHDCRHKEDAGVICSEPT 1139
DB 1148 GKAIATVTRAPRGQGPILYLDNINCTGSETSLNCSHSLGTHDCHDSEDAGVICSQPD 1207
QY 1140 ALRL---YSETETESACRLEVFYNGTWGSGVRRNITTAIAGIVCQRLGCGGNGVVS LAP 1196
DB 1208 TIRLAGRSKYE-----GRVEILQNGAMGTVCDDNDWHDAAVVCRELGYATLTLEATSOA 1262
QY 1197 LSKTSGFWVDVDDIOCPKTHISIWOLCSAPWERRISSPAETWITCEDRIRVR--GGDTE 1254
DB 1263 SPQGTGTAILDDLRCSGREIRLVDCPNGLQHNCQHLLEDAGVVCQDVIEURLVDGQTP 1322
QY 1255 CSGREIWHAGSWGTVCCDSDWDLAEAVVCCKQLGCGSALALRADASFGQGTGTIWLDDMR 1314
DB 1323 NAGRVELRYQOGMGVTCDDNGWDVDAVVCRELGPRAIANNQAFYGPICGTVLLSDIS 1382
QY 1315 CKGNESFLWDCHAKPWGSDCKHEDAGVRC-----SQOSLKSLSNASSGHIALILSSITF 1368
DB 1383 CGGQBATLFCPSNPGVGLNSCHNEDAGVVCMPKVRLSG-----SNNPNEGLEVFMNGLM 1438
QY 1369 G 1369
DB 1439 G 1439
RESULT 7
T18524
scavenger receptor cysteine-rich protein homolog srcrm2 - Geodia cydonium
C:Species: Geodia cydonium
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
C:Accession: T18524
R:Blumbach, B.; Pancer, Z.; Diehl-Seifert, B.; Steffen, R.; Muenkner, J.; Mueller, I.
J. Cell Sci. 111, 2635-2644, 1998
A:Title: The putative sponge aggregation receptor: Isolation and characterisation of
A:Reference number: 218947; MUID:98369060; PMID:9701562
A:Accession: T18524
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2043 <BLU>
A:Cross-references: EMBL:Y14953; NID:el364818; PID:el364819; PIDN:CAA75175.1
C:Genetics:
A:Gene: SRCRM2
Query Match 27.6%; Score 2246.5; DB 2; Length 2043;
Best Local Similarity 32.4%; Pred. No. 1.4e-137;
Matches 514; Conservative 196; Mismatches 588; Indels 287; Gaps 31;
QY 7 SHWIDEGRCCHQNLFSAVVTCILLNSCFLLISFNCTGTLRLV-NGDPCSGTVEYKF 65
DB 114 AWH-SVGSCHQNNSGVVCT-----DDSAVRLVEGDMNKRGRVEYH 156

```
Qy 66 QGONGTVCDDGNTTASTVVCVKQLGCPFSFAMFREGQAVTRHGKIMLWDDVSCYNESALW 125
Db 157 ADEGSCICDDYFGGEAVVICHHELGFSAEDSHGGTFFPOGDGLIWMNDLQCTCHESSIA 216
Qy 126 ECQREHWSHNCYHGEDVGVNCEAN-----LGLRLVDGNNSCS-----GRVEVKFQ 173
Db 217 ECTFPFGGEHNCQREDAGVCTPATDHPFYIPIRLNNGTSSNATHGLOYGRVEFIN 276
Qy 174 ERMTICDDGNLNTAAVVCRCPCSPSSISSGVNPAVLN---PTWLDDILCOGNELA 230
Db 277 NTWGTVCDDGNGIEDATACRLQFYGALLASSRSEVAQVAPSPYIWDVVLCHGDEVA 336
Qy 231 LWNCRHGWGNDCHSHNEDVTLTCYDSDLE-----LRLVGGTNRCM-----GRVEL 277
Db 337 LSECHHRTTGVHDCSHVEDAGVFC--TTNDIQAPAKVRLVRNNGNSVDEDCGHVEGRVEV 395
Qy 278 KIQRWGTVCCHKNNNAADVVCVKQLGCGGTALHFAGLPHLQSGSDVWLDGVSCSNGNEF 337
Db 396 NFYGVWGTICDDYWSLADGVICKMLGFTYADEVGYSGAFGIGTGPIMWDDVQCDGTET 455
Qy 338 LWCDRHSGVTNFDCLHONDVSVICS---DGADLELRADGNNSGRVEIRIHEQWWTIC 394
Db 456 IDHCDFFGHWGNYCHYEDASVVCYRQPGSD--VRLVGGSSMEGRVEVLNGVWGTVC 513
Qy 395 DONKNNEOALVVCVKQLGCPFSFVSGRRAP-----SNEARDIWNISICTGNESALWDC 448
Db 514 DDSGAEDAQVICHQLG----FYGTATAVRGTQGFASSSQPIWLDDVGCTGSEMYLSDC 569
Qy 449 TYDGKAKRTCFRRSDAGVIC-----468
Db 570 SSSGFGNNHNCYHEDAGVGVCOGTSAFIPVRLNSGTRNPGTVEIFYQNRWGTICDIFWTL 629
Qy 469 -----468
Db 630 YDANVVCRLQYDGNAYNATHSSYCGSLTTVATYDCFGQESQLANCTGFRYPPIPSW 689
Qy 469 -----SDKADLRLVGAHSPCYGRLEVYQGEWGTVCHDRWSTNAAVVC 514
Db 690 YCGDNTVAGVMCIGTEDGPAPIRLVGSEASEGRVEILVQGWGTICDDIWDQADAEEVC 749
Qy 515 KOLG-CGKPMHVFMTYFKEASGPTWLDVSCIGNESINWDCSHGKGNKNCVHREDVIV 573
Db 750 RLQYLSEGATALQFAHFGEGSGPIVLDVNCCTGLELYTDCPSGFGYIHNCAHSEADGV 809
Qy 574 TCSG-----DATWGLRVGGS---NRCSGRLEVYFQGRWGTVCDDGWNKAAVVCSSOLD 625
Db 810 RCNARGPQTQYTPIRLVRNGTVITONBGTVEILHNGTWSAVCDDYWCYTEAVVACHML- 868
Qy 626 CPSSIIIGMLGNASTGY-----GKIWLDDVSCDGEDSLWSCRNSGNGNDCSHSE 676
Db 869 -----GFATAVRAVTRSPHGAVDGDTFLDNVRCACGTERTIEDCYHSSYSRNCDRTO 920
Qy 677 DVGVTCSO--ASDMELRLVGGSSRCAGKVEYNVOGAVGILCANGWGMNIAEVVCRLGECG 734
Db 921 QAGVACTNYTASEYPIRLVGGSGPHEGRVEIYOGVWGTVCDDGWSQPDADVVCRLQGYA 980
Qy 735 SAIRVSREPHEPTERTLHILMSNGCTGGEASLWDCIRNEWQOTACHLNEASLICSASRHQ 794
Db 981 NASRATVRAEFGRTGETIWLNDVACTGFENSLDECNSGWDHNCCHREDAGAVCOGELL 1040
Qy 795 P-RLVGADMPSCGRVEVKHATVRSVCDSDFSLHAANYLCRELNCGDALISVGDHFGKG 853
Db 1041 PIRLRDGSNLEGRVEVFNKWTGTVCDDFDWLDGATVVCRLQGYPEAVRAPEARFCSG 1100
Qy 854 NGLTWAKEFCQEGSETHALCPVQHPEDTCIHSREVGVVC-----SRYTDVRLVNGKS 907
Db 1101 EGPWLDDVAVCGTETSLFESSGSLGQHNCOHFDAGAVECNDHITPSPVTDLRAGGTN 1160
Qy 908 QCDGOVEINVLHGWSGLCDTWDPEADARVLCRLSCGTALTGTGGKI--GERSVRVWGH 965
Db 1161 ATEGRVEIFFNNTWGTICDDSDWIDRAEVVCRYLGFESAIALSNGYFCAGDPOPIWLD 1220
Qy 966 RFHCLGNESSLNDNQMTVLGAPPCIHGNTVSVICTGSLTQPLFPLCLANVSDPYLSAVPEG 1025
```

```
Db 1221 DVDCFGSSEITITSCLTSALEHNCARHVEDAGVRC--YISHPL-----1260
Qy 1026 SALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVVCOKLGCQVAFNA 1085
Db 1261 -----LRLVNGNTEGGRVEVHRGWSGTVCDDHWEIDANIVCKELGFARALISA 1310
Qy 1086 TVSAHFEGEGSGPIWLDDLNCITGTESHLMQCPSRGQWQDCHRKEDAGVICSEF----TAL 1141
Db 1311 SGFATFEGEGSGTIWLDDVQCDGTEATIFNCFSSGNGNHNCLHSEDAISAICTAIPVNPV 1370
Qy 1142 RLYSETETESACGLEVFYNGTWSVGRNRNITTAIAGTVLCOLGC--GENGVVSLAPLSKT 1200
Db 1371 QLIDGPSVHE--GRLOIYNNNEGTVCDDQWGYDEANYVCKSLGPGADPNLSLLNSYGP 1428
Qy 1201 GSGFMWDDDIQCPKTHISIWOLCLAPWERRISSPAEETWITCEDR---IRVRGGDTECSG 1257
Db 1429 GSGHILWDDVDRCRDEFFIQDNCNADIGENNCGHYEDVGLKCLPNTLVDRLMNGNTSAG 1488
Qy 1258 RVEIWHAG--SWGTVCCDSDWDLAAEAVVCQQLGCGSALAAALRADSAFSGQGTGTIWLDDMRCK 1316
Db 1489 RVEVNYNGEENGTVCCDNDWDIKDADVVCRLMDFKSAQAPREAFPGDGLGVWEDFELCS 1548
Qy 1317 GNEFLWDCCHAKPGQSDCGHKEKEDAGVRCQSLKSLNASSGHLALILSSIFGLLLVLF 1376
Db 1549 GYEDSLLECSHAGVKVHNCRHSEDAVSVC-----1578
Qy 1377 ILEFTWCRVQKQKHLPLRVSTRRRG 1401
Db 1579 --MLTVCRDFSQEHGTFOISNRQFG 1601

RESULT 8
T30549
hensin - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: T30549
R;Takito, J.; Hikita, C.; Al-Awqati, Q.
J. Clin. Invest. 98, 2324-2331, 1996
A;Title: Hensin, a new collecting duct protein involved in the in vitro plasticity of
A;Reference number: 220851; MUID: 97096804; PMID: 8941650
A;Accession: T30549
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1594 <TAK>
A;Cross-references: EMBL:AF043112; NID: 94105083; PID: 94105084; PIDN: AAD02242.1

Query Match 26.5%; Score 2156.5; DB 2; Length 1594;
Best Local Similarity 36.6%; Pred. No. 7.1e-132;
Matches 475; Conservative 132; Mismatches 387; Indels 303; Gaps 29;

Qy 46 LELRLVNGDGPCSGTVEVKFQOGNGTVCDDGNTTASTVVCVKQLGCPFSFA---NFRFGQ 102
Db 51 LELRLANGGRCQGRVEVLYQSGWGTVCDDGWDINDAQVVCRLQCGGMVAVSAPGSAREFQ 110
Qy 103 AVTRHGKTLWDDVSCYGNESALWECQHRWESHGSHNCYHGEDVGVNCEY-----149
Db 111 G---PGQIVLDDVSCSGQEPYLMWSCHHRGWSLHNCGHQEDAGVICSDAMWTPSPPTDTP 167
Qy 150 -----EANLRLVLDGNNCSGRVEVKFQERWGTICDDGWNLTAA 190
Db 168 TTVIYESTPHFPGLLELVFPPSGLELRLANGSDRCQGRVEVLYQSGWGTVCDDGWDINDAQ 227
Qy 191 VVCRLGCPSPSFSSGVNPAVLN-----PIWLDDILCOGNELANCRHGRGNHDC 244
Db 228 VVCRLQCGMA-----VSAPGSARFQGGQIWLDDVSCSQEPYLMWSCHHRGWSLHNC 281
Qy 245 SHNEDVTLCYDS-----SDLELRLVGGTNRCMGRVVELKIQ 280
Db 282 QHEDAGVICSDAVPTTTPPDTPVTIYESSVPFSGLELRLANGSDRCQGRVEVLYQ 341
Qy 281 GRWGTVCHHKWNNAADVVCVKQLGCGGTALHFAGLPHLQSGSDVWLDGVSCSNGNESFLWD 340
```

Db 342 GSWGTVCDDGWDINDAQVVCRLGCGTAVSAPGSARFGQCGQVQLVDDVSCGQEPYLS 401
QY 341 CRHSTVNFDCIHLQNDVSVICSDG-----ADLELRADGDN 376
Db 402 CHRWGLSHNCGHQBAGVICSADAMATPPLDTPWTTIHESTVPFPGGLEQLANGSD 461
QY 377 NCSGVEVRIHEQWTTICDQWKNKEQALVCKQLGCPFSVFGSRRAKPSNEARDIWIINSI 436
Db 462 RCQGRVEVLYQ---CTVCDDGWDINDAQVVCRLGCGMAVSAPGSARFGQVQLVDDV 518
QY 437 SCTGNESALMDCTYDGAKECTFRSDAGVICS-----KA 472
Db 519 SCGQEPYLSCHRRGLSHNCGHQBAGVICSADAMATPPLDTPWTTIHESTVPFPGGLEQLANGSD 578
QY 473 DLRLVGAHSCYCRLEVKVQGEWGTVCHDRWSTNNAAVVCKQLGCKPMHVFQMTVEK 532
Db 579 GLELRANGSDRCQGRVEVLYQSGWGTVCDDGWDINDAQVVCRLGCGTAVSAPGSARFG 638
QY 533 EASGPIWLDVSCIGNESINIMDCSHGKNCVHREDVIVTCSG-----DATW--- 581
Db 639 QGPGQIVLDDVSCGQEPYLSCHRRGLSHNCGHQBAGVICSADAMATPPLDTPWTTIHESTVPFPGGLEQLANGSD 698
QY 582 -----GLRLVGGNRCGRLEVYFGRWGTVCDDGWSKAAAVVCSQDPCPS 628
Db 699 VIYESTPVHISGLQLRLVNGSDRCGRVEVLYQSGWGTVCDDGWDINDAQVVCRLGCGTAVSAPGSARFG 758
QY 629 STIIGMGLGNASTGYCKIWLDDVSCDGDSDLSWNSGNGNDCSHSDVGVICSDA--SD 687
Db 759 ASAPASQAQFGQSSGIVLDDVSCGSEPNLWSCSHRGWLSHNCGHEDAGVVCSGPDSR 818
QY 688 MELRLVGGSSRCAGRVENVOGAVILCANGMNAIEVVCRLGCGSAIRVSRPHETE 747
Db 819 LAVRLVNGSTRCGRVLEVYRGSWGTVCDDGWDINDASVVCRLGCGMAVSAPGSARFG 878
QY 748 RTLHILMSNGTGEASLWDCIRNEWQOTACHLWNASLICSARHQPRLVGCADMPCSR 807
Db 879 GSGSIFLDEVSCGQEPYLSCHRRGLSHNCGHQBAGVICS----- 921
QY 808 VEVKHADTWRSVC---DSD--FSLHAANVLCRELNGDAISLSVGDHFGKGNLTWAEKF 862
Db 922 -----DGWTTVPAPTDDWEPVITTVGSSNCGGFLYNATGSFSP----- 965
QY 863 QCEGSETHIALCPIVOHPEDTCIHSREVG-----VCSRYTDVRLVNGKSC----- 909
Db 966 SYPGYVNNALC-----VMEIAPVSGYLINLFGSOLRL--EQHSYCNFDYVEIF 1012
QY 910 DQGVETNLGHWSLCTDHWDPEDARVLCRLSCGTALSTTGKYGIGERSVRVWGHRRFC 969
Db 1013 DGSTSSLL---GKICN-----DSGOIFITSSNR----- 1038
QY 970 LGNESILLNCOMTVLGAAPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALI 1029
Db 1039 -----MTVL-----FRSDISVQNTGFLA-----WYNSFP----- 1062
QY 1030 CLEDLRLVLDGDS---RCAGREVIYHDFWGTICDDGWDLSDAHVVCKLCCGVAFNAT 1086
Db 1063 --RDASLRVLSGSSYGACAGREVIYHGGRWGTVCDDGWDINDAQVVCRLGCGMAVSAP 1120
QY 1087 VSAHFCEGSGPTIWLDDNCTGTESHLCQPSRGWGWDCRKHEDAGVIC-----SEFTA 1140
Db 1121 GGAIFGSGGPTIWLDDNCTGTESHLCQPSRGWGWDCRKHEDAGVIC-----SEFTA 1180
QY 1141 LRLYSETETESACGLR-----EVFYNGTWGSGVR 1169
Db 1181 PNISTSNASVSCGFFLSQHSRGFSFPFPGNYPNNAR 1217

RESULT 9

T42721

CRP-ductin-alpha precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 23-Mar-2001

C:Accession: T42721
R:Cheng, H.; Bjerknes, M.; Chen, H.
Anat. Rec. 244, 327-343, 1996
A:Title: CRP-ductin: a gene expressed in intestinal crypts and in pancreatic and hepa
A:Reference number: 22241; MUID:96362470; PMID:8742698
A:Accession: T42721
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2083 <CHE>
A:Cross-references: EMBL:U37438; NID:q1276646; PID:q1276647; PIDN:AAC52505.1
A:Experimental source: strain BALB/c; jejunal epithelial cells
C:Keywords: transmembrane protein
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-2083/Product: CRP-ductin-alpha #status predicted <MAY>

Query Match 24.1%; Score 1962.5; DB 2; Length 2083;

Best Local Similarity 28.9%; Pred. No. 3.7e-119;

Matches 49; Conservative 164; Mismatches 503; Indels 557; Gaps 41;

QY 42 NGTDD--LELRVNGDPCSGTVEVKFGQMGTCVDDGWNNTASTVVCVKQLGCPFSFAM-- 97

Db 29 DGTDSGLAVRLVNGDRCQGRVEILYQSGWGTVCDDGWDINDANVVCRLGCGLAVSAPG 88

QY 98 -FRGQAVTRHCKIWLDDVSCYGNESALWECQHRMGSHNCHYHGEDVGVNCGV----- 149

Db 89 NARFGQ---SGPIVMDVACGGYEDYLWRCSHRGWLSHNCGHQEDAGVICSDSQTSPT 145

QY 150 -----EALGLRLVDGNNSCSGRVEVKFQE 174

Db 146 PQWNPNGTNDVFPYPTQTAEQTTIPDTPIGTDSGLAVRLVNGDRCQGRVEILYQ. 205

QY 175 RWGTICDDGWNINLNTAAVVCRLG-----CPSEFTSSGVVNSPAVLRLWLDLILCO 225

Db 206 SWGTVCDDSDVSDANVVCRLGRLWGLVSAQEMPFGGQGLV-----LDDVACG 255

QY 226 GNELALNCRHRCGNHDCSHNEDVTLACYDS----- 257

Db 256 GYENTLWCSHQGLSHNCGHQBAGVICSASQSSPTPGWNPNGTNDVFPYPTQTAE 315

QY 258 ---SDLELRVNGTNRCMGRVBLKIQGRWGTVCCHRWKNAADVVCVKQLGCGTALHFAGL 314

Db 316 GTDGLAVRLVNGDRCQGRVEILYQSGWGTVCDDGWDINDANVVCRLGCGMAVSAPGN 375

QY 315 PHLQSGSDVWLDDVSCSGNESFLWDCRHSHTVNFDCIHLQNDVSVICSD----- 363

Db 376 AYGFGSGSIVLDDVACTGHEDYLWRCSHRGWLSHNCGHEDAGVICSASQSSPTPDVF 435

QY 364 -----GAD--LELRADGNSCSGRVEVRIHEQWTTICDQWKNKEQ 402

Db 436 YPTDQTAEQTTVPDTPIGTDSGLAVRLVNGDRCQGRVEILYQSGWGTVCDDGWDIND 495

QY 403 ALVVCVKQLGCPFSVFGSRRAKPSNEARDIWIINSICTGNESALMDCTYDGAKECTFRRS 462

Db 496 ANVVCRLGCGLAVSAPGSARFGQGTGPVMDVACGGYEDYLWRCSHRGWLSHNCGH 555

QY 463 DAGVICS-----DKADLDLRLVGAHSPCVGRLE 490

Db 556 DAGVICSASQSSPTPDVFPYPTQTAEQTTVPDTPIGTDSGLAVRLVNGDRCQGRVE 615

QY 491 VYQGEWGTVCCHDRWSTNNAAVVCKQLGCGKPMHVFQMTYFKAESGPIWLDVSVICGNS 550

Db 616 ILYQSGWGTVCDDGWDINDANVVCRLGCGLAVSAPGSARFGQGTGPVMDVACGGYED 675

QY 551 NIWDCHEHSGWGHKNCVHREDVIVTCSG-----D 578

Db 676 YLWRCSHRGWLSHNCGHQBAGVICSASQSSPTPDVFPYPTQTAEQTTVPDTPYTTIGTE 735

QY 579 ATWGLRLVCGSNRCGRLEVYFGRWGTVCDDGWSKAAAVVCSQDPCFSSIIIGMGLGNA 638

Db 736 NSLAVRLVNGDRCQGRVEILYQSGWGTVCDDGWDINDANVVCRLGCGLAVSAPGSARF 795

QY 639 STGYGKIWLDDVSCDGDSDLSWNSGNGNDCSHSDVGVICSDA----- 685

Db 796 GQGTPIVMDVACGGYEDYLWRCSHRWLSHNCGHEDAGVICSASQSSSTPDVFYPT 855
QY 686 -----SDMLRLVGGSSRCAGKVEVNVQAGVITLCANGMGMTAEV 726
Db 856 DQTVTEOTVTPDTPGTENSLAVRLENGDRCGRVEILYQSGWTVCDSDWTKDANV 915
QY 727 VCRQLCGSAIRVSRPHEPFTLHLILMSNGSTGGEASLWDCIRWEWKOTACHLNMES 786
Db 916 VCRQLCGMAVSAFNAYFGPGSGVILDDVACTGHEDYLWCSHRGLSHNCGHEDAG 975
QY 787 LIC-----AHROPRL-----VGADMPGSG----- 806
Db 976 VICSDAQIGSTTRPDLPMTPTTPTTELTTPYDFDWTMTTSDYSCGGLLTQPSGQFSS 1035
QY 807 -----RVEYKHADTW-----RSVC 820
Db 1036 PYPSPNPNARCWKIVLPNMNRVTVFVTDVQLEGSCNVDYILVYDGPENSSLIARVC 1095
QY 821 DSD-----FSLHAANVLCRELNGDALSLSVGDHFGKNG 855
Db 1096 DGSNGSFTSTGNFMSVVFITDGSVTRRGFOAHYVSTVSTNYSCGGLLTQPSGQFSSPYYP 1155
QY 856 LTWAERFQCGSETHLALCPVQHPEDCTHSREVGVVCSRYTDLVRLVNCQKQ----- 908
Db 1156 SNTPNNARCWE-----ILVP-----FDTVQLEGSCNVDYILVYD 1198
QY 909 -----CDGQVEINVLHGWSLCTDHPEDARVLCRLQSLCTALSTTGKRYIG 956
Db 1199 GPQYNSSLIARVCDGSN-----GSFTST-----GNFMSVV---FIT 1231
QY 957 ERSVRVWGHFRHCLGNESLNDNCOMTVLGAAPCIHGNTVSVICTGSLTOP----- 1006
Db 1232 DGSVTRRGFOAHYYS-----TVSTTPPVPIPTDDYSCGGLLTLPSCGQFSSPHY 1281
QY 1007 -----LFPCLANVSDPYLSAVPEGSALICLEDKRLVLDGDSRCAGRVEIYH 1053
Db 1282 SNTPNNARCWEILVPMNRVTVAFTDVQLEGG---CNYDIYL-VYDQPE-----YN 1329
QY 1054 DGFMTGICDDGDLSDAHVVCQKLGCVAFNATVSAHFGEVGPVILDDLNCST---GPESH 1111
Db 1330 SLLIARVCD-----GSGNSTST-----GNFMSVVFITDGSVTRRGFOAH 1369
QY 1112 LMQCPSRGW-----GQHDCHR---KEDAGVICS----- 1136
Db 1370 YYSTVSTNYSCGGLLTQPSGQFSSPHYSPNPNVRCSEWILVPSMNRVTVAFTDVQLEG 1429
QY 1137 --EFTALRLYSETEES-----CAGLEVFYNGTWGSGVRRNTTALAGIVCRQLCCGE 1188
Db 1430 GCSFDYILVYDGPENSSLIAPVCDG-----FNGSFTSTG---NFMS----- 1468
QY 1189 NGVYVSLAPLSKTSQGFMMVDDIQCPKTHSIWQCLSAPEWERRISSPAETWITCED---R 1245
Db 1469 --VFTDGSVTRGF-----QAYYSTVSTSLHFNITGNDSLSA 1507
QY 1246 IRRVGGDETCGRVEIWHAGSWGTVCDSDWDLAEVYVCOQLCGSALALRDAFSQGT 1305
Db 1508 LRLVNGSNRCRGEVILYRSGWTVCTVCDSDWISDANVVCRLCGSALSAPGNWAFQGS 1567
QY 1306 GTWLDDMRCKNGESFLWDCAPWGQSDCGHEDAGVICS 1346
Db 1568 GLVILDDVSCSGYESHLWNCHHPGWLHNCRHSEDAGVICA 1608

RESULT 10

S56744

mucin (clone pgM7-1) - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C:Date: 27-Oct-1995 #sequence_revision 30-Jan-1998 #text_change 20-Apr-2000

C:Accession: S56744

R:Nunes, D.P.; Keates, A.C.; Afdhal, N.H.; Offner, G.D.

Biochem. J. 310, 41-48, 1995

A:Title: Bovine gall-bladder mucin contains two distinct tandem repeating sequences: ev

A:Reference number: S56744; MUID:95374471; PMID:7646470

A:Accession: S56744

A:Molecule type: mRNA

A:Residues: 1-600 <NUN>

A:Cross-references: GB:S78981; NID:g1042036; PIDN:AAB35069.1; PID:g1042037

F:1-102/Domain: scavenger receptor cysteine-rich domain homology

F:126-229/Domain: scavenger receptor cysteine-rich domain homology (fragment) <SRC1>

F:253-356/Domain: scavenger receptor cysteine-rich domain homology <SRC2>

F:380-483/Domain: scavenger receptor cysteine-rich domain homology <SRC4>

Query Match 13.2%; Score 1076.5; DB 2; Length 600;

Best Local Similarity 38.5%; Pred. No. 2.9e-62;

Matches 223; Conservative 52; Mismatches 157; Indels 147; Gaps 13;

QY 48 LRLVNGDGPSCGTVKVFQGWGTVCDGWNTTASTVVKQLGCPFPFAM---FRFGQAV 104

Db 2 LRLVNGSDRCQGRVEVLYGGWGTVCDDSDWTDNDANVVCRLGCGWALSAPGDARFGG- 60

QY 105 TRHGKIWLDDVSCYGNESALWECQHRWGSNHCYHGEDVGVNC----- 147

Db 61 --SGPIVLDVCGSGYETYLWCSHSPWNTNHCGHSEDAVICSASQTQSTVVPDWLYPT 118

QY 148 --YG-EANLGLRLVDGNNSCSGRVEVKEFQERWGTICDDGNLNTAAVVCRLGCPSPFIS 204

Db 119 TDGTESEGLALRLVNGDRCQGRVEVLYRGSWGTVCDDSDWTDNSNVVVCRLGCG- 172

QY 205 SGVYNS-PAVLR-----PIWLDDILCOGNELALWNCRHRCGNHDCSHNEDVTLC--- 254

Db 173 -GWANSAPGNARFCQGGSPVILDDVCGSGYETYLWCSHSPWNTNHCGHSEDAVICSAS 231

QY 255 -----YDSSD-----LELRLVGGTNRCMGRVELKIQGRWGTVCVCHHKWNNAAAD 297

Db 232 QTQSTVVPDLWYPTDYTESGLALRLVNGSDRCQGRVEVLYRGSWGTVCDDSDWTDNDAN 291

QY 298 VVKQLCGGTALHFAGLPHIQSGSDVVLVDGVSCHGNEFLWDCRHSGTVNFDCLHQNDV 357

Db 292 VVCRLGCGMGISAPGDARFGQGGPIVLDVCGSGYETYLWCSHSPWNTNHCGHSEDA 351

QY 358 SVICSADGADLELRLADGNSNCSGRVEVRIHQWNTICDQNKNEQALVVCQLGCPSPVF 417

Db 352 SVICSA-----SQTQSTV----- 365

QY 418 GSRRAPSNREARDIWIINSICTGNESALWDCYDGKAKRTCFRRSDAGVICSADKADLDR 477

Db 366 -----PDLW-----YPTDYGCT-----ESGLALR 384

QY 478 LVGAHSPCYGRLEVKYQGEWGTVCCHDRWSTRNAAVVCQLGCGKPMHVFQMTYKEASGP 537

Db 385 LVNGSDRCQGRVEVLYGGWGTVCDDSDWTDNDANVVCRLGCGSISAPGDARFGQGGP 444

QY 538 IWLDDVSCIGNESNIWDCHEHSGKGNCHVHREDVIVTCS 576

Db 445 ILLDDVCGSGYETYLWCSHSPWNTNHCGHSEDAVICS 483

RESULT 11

A57190

ebnerin precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999

C:Accession: A57190

R:Li, X.J.; Snyder, S.H.

J. Biol. Chem. 270, 17674-17679, 1995

A:Title: Molecular cloning of Ebnerin, a von Ebner's gland protein associated with ta

A:Reference number: A57190; MUID:95353552; PMID:7629065

A:Accession: A57190

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A:Molecule type: mRNA

A:Residues: 1-1290 <LIA>

C:Cross-references: GB:032681; NID:g975346; PIDN:AAC52248.1; PID:g975347

C:Superfamily: scavenger receptor cysteine-rich domain homology; C1r/C1s repeat homol

C:Keywords: extracellular protein; transmembrane protein

F:1-77/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SRC

F:142-245/Domain: scavenger receptor cysteine-rich domain homology <SRC1>
F:289-392/Domain: scavenger receptor cysteine-rich domain homology <SRC2>
F:583-689/Domain: C1r/C1s repeat homology <C1R1>
F:713-816/Domain: scavenger receptor cysteine-rich domain homology <SRC3>
F:838-944/Domain: C1r/C1s repeat homology <C1R2>

Query Match 12.98; Score 1048.5; DB 2; Length 1290;
Best Local Similarity 29.78; Pred. No. 4.5e-60;
Matches 289; Conservative 114; Mismatches 290; Indels 281; Gaps 34;

Qy 286 VCHKWNAAADVCKQLGCGCTALHFAGLPHLQSGSDVWLDGVSCSGNESFLWDCRHS 345
Db 1 MCDSDWINDANVCRQLGCGWALSAPGSAQFGGLGPIVLDVACRGEAYLWSCSRG 60
Qy 346 TVNFCDLQNDVSVICSDGADLELRLADGNNCSGRVEVR:HEQWTTICDQWKNQALV 405
Db 61 WLSHNCGHQEDAGVICS- - - - -SQTSSPTG- - - - -WW- - -NPGTNDVIY 100
Qy 406 VKQLGCGFVSRRAPKPSNEARDIWNISCTGNESALWDCYDYGAKKTCFRSDAG 465
Db 101 DQET- - - - -TETSSTSTPDW- - - - -NHGTTINDVIYD- - -TQETEGTDSG 142
Qy 466 VICSDKADLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRAAVVCKQLGCGKPMHV 525
Db 143 - - - - -LAVLVNGDRGRVEILYQSGWIVCDSDWINDANVCRQLGCGWALS 194
Qy 526 FGMYFKASGPIWLDDVSCIGNESNIWDCSHGSKHNCVHREDIVITCS- - - - - 576
Db 195 PGSAGFGGSGSIVLDDVACRGEAYLWSCSRHGLWLNCHQEDAGVICSQTSSTPT 254
Qy 577 - - - - -GDATWG- - - - -LRLVGGNRCGRLEVFQGRWG 605
Db 255 DSQTSSTPTPGMNPGGTNDVSYGPEQTTDATDGLAVLVNGDRGRVEILYQSGW 314
Qy 606 TVCDDGWSKAAAVVCSLDPCSSIIIGLGNASTGYKIKLDDVSCDGEDSDLSWCSRNS 665
Db 315 TVCDDSDTKDANVCRQLVCGWALSAPGSAHFGGSGSIVLDDVACRGEAYLWSCSR 374
Qy 666 GWNNDCHSDEVDGVICS- - - - -DM- - - - -ELRLVGGSRCAKGVVNVQGVAG 712
Db 375 GWLSHNCGHEDAGVICS- - - - -CGLLTLTPYGFSSPYPGS- - - - -YPNNARC- - - - -LW- 474
Qy 713 ILCANGWGNIAEVVCRQLGCGSALRVS- - - - -REPHEFERTLHILMSNGCTGGEASLWD 768
Db 434 - - - - -ADWTTTPPEYT- - - - -CGLLTLTPYGFSSPYPGS- - - - -YPNNARC- - - - -LW- 474
Qy 769 CIRWEKQATACHLNEASLICS- - - - -SAHRQLVADMP- - - - -PCSRVEYKADTWRSVCDSDFL- - - 826
Db 475 - - - - -KIFYSSMNRVTVTDVQLEGG- - - - -CNYDYILVF 505
Qy 827 - - - - -HAANVLCRELNGDAISLVGDHFGK- - - - -NGLTWAEKFOCE- - - - - 865
Db 506 DGPENNSLIAR- - - - -VCDGFGSFTSTQNFMSVFIIDGSGVTRRGFOADYYS 554
Qy 866 - - - - -GSETHLALCPV- - - - -QHPEDTC- - - - -IHSREVGVCSRYTDLVNLVCKSCDQGV 913
Db 555 TPISTSTSTPTFPVTDWTTTPPEYTCGGLTLTPYGFSSPY- - - - - 598
Qy 914 EINVLHGWSGLDTHWDPEDARVLCR- - - - -QLSCGT- - - - -ALSTTGK 953
Db 599 - - - - -YPGSY- - - - -PNNARLWKIFVPSMNRVTVTDVQLEGGCNYDYILGFDGPE 646
Qy 954 YIGERSVRVWGHFRHCLGNESLLDNQMTVLGAPPCIGHNTVSV- - - - -ICTGSLTQPLF- - - 1008
Db 647 YNSLIARV- - - - -CDGNSGFTSTQ- - - - -NFMVSVFITDGSVTRRGFOAD 688
Qy 1009 - - - - -PCLANVSPDYLISAVPEG- - - - -SALICLEDKRLVLDGSRCAGRVEIYHDFGWTICD 1062
Db 689 YVSTPIRSTPTPTPTPTPTITGNDSSLV- - - - -LRLVNGTNCGRVEILYRGSWVPCAD 742
Qy 1063 DGDWLSDAHVVCQKLGCGVAFNATYSAHFEGSGPIWLDDNLCTGTESHLWQCPGRGWQ 1122
Db 743 DSWDINDANVCRQLGCGSALSAPGNWFGGSGGLVLDVYSCSGYESHLLNCRHPGLV 802

Qy 1123 HDCHRKEDAGVICS 1136
Db 803 HNCRHVEDAGVICS 816

RESULT 12
TI7405
C:Species: Strongylocentrotus purpuratus (purple urchin)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
C:Accession: TI7405
R:Pancer, Z.; Rast, J.P.; Davidson, E.H.
Immunogenetics 49, 773-786, 1999
A:Title: Origins of immunity: transcription factors and homologs of effector genes of C:Species: Strongylocentrotus purpuratus (purple urchin)
A:Reference number: Z18253; MUID:99328904; PMID:10398804
A:Accession: TI7405
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1036 <PAN>
A:Cross-references: EMBL:AF076513; NID:g4164530; PID:g4164531; PIDN:AA05493.1

Query Match 10.6%; Score 866.5; DB 2; Length 1036;
Best Local Similarity 27.18; Pred. No. 2.3e-48;
Matches 229; Conservative 80; Mismatches 227; Indels 309; Gaps 18;

Qy 312 AGLPHLQSGSDVWLDGVSCSGNESFLWDCRHSHTVNFDCIHLQNDVSVICSDGADLELRL 371
Db 478 SGIPAL- - - - -LDNIAADVDSIL- - - - -SVLTDA- - - - -QOCSTIVQEGDLALTY 519
Qy 372 ADGSNNCSGRVEVR:HEQWTTICDQWKNQALV- - - - -CKQLGCFPSVFGRR 421
Db 520 IPGLD- - - - -PYALNLQIYQITGMNGTNQLCQVVADAFQPGDHHCP- - - - -GMTFECGSGC 572
Qy 422 AKPSNE- - - - -ARDI- - - - -WINSISCTGNESALWDCYDYGAKKTCFRSDAGVIC- - - - -SDKADL 474
Db 573 GPGSCDNLVPRDIPCLFCFVGCFCPEGLVKD- - - - -QDGDRCIPVDQCG 618
Qy 475 DLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRAAVVCKQLG- - - - -CGKPMHVGMTY 530
Db 619 AIRLVGNEAEGRVEIYQNGVMTGICDSSWGITDASVVCRLMGFGQAGSAP- - - - -GSAH 674
Qy 531 FKEASGPIWLDDVSCIGNESNIWDCSHGSKHNCVHREDIVITCGDATWLRVLVGSN 590
Db 675 FQGGTPIQLDDVCGTGAETIFDCAHPAFGVHNCACHEDAGVFCI- - - - -ASQDVRVLVGSN 732
Qy 591 RCSGRLEVYFQGRWGTCDGWNKAAAVVCSQLDPCSSIIIGMGLGNASTGYKIKLDDV 650
Db 733 EAEGRVEIYQNGVMTGICDSSWGITDASVVCRLMGFGQAGSAPGSAQFGQGTGLIQLDDV 792
Qy 651 SCDDGSDLSWCRNGNCSHSDVGVICSDASDMLRLVGGSSRCAGKGVNVQGA 710
Db 793 GCTGAETIFDCAHPAFGVHNCACHEDAGVFCI- - - - -VRLVGGSEAEGRVEIYQNGV 850
Qy 711 VGILCANGWGNIAEVVCRQLGCGSALRVSREPHEFERTLHILMSNGCTGGEASLWDCI 770
Db 851 WGTICDSSWGITDANVCRQLMGFGQAGSAPGSAHFGQGTGPIQLDDVCGTGVQETIFDC- 909
Qy 771 RWEWKQACHLNEASLICS- - - - -SAHRQLVADMP- - - - -PCSRVEYKADTWRSVCDSDFLHAAN 830
Db 910 - - - - -AH- - - - -PPFGVH- - - - - 917
Qy 831 VLCRELNGDAISLVGDHFGKNGLTWAEKFOCESGSETHLALCPVQHPETC:HSREV 890
Db 918 - - - - -NCAHYEDA 925
Qy 891 GVCSRYTDLVNLVCKSCDQGVINVLHGWSGLDTHWDPEDARVLCRQLSCGTALSTT 950
Db 926 GVVCILVSDV- - - - - 935
Qy 951 GSKYIGERSVRVWGHFRHCLGNESLLDNQMTVLGAPPCIGHNTVSVICTGSLTQPLFP 1010
Db 936 - - - - - 935

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 05:30:12 ; Search time 21 Seconds
(without alignments)
2869.767 Million cell updates/sec

Title: US-09-759-130B-381
Perfect score: 8138
Sequence: 1 MMLPQNSWHIDFGRCCHQ.....CEDASDTSLGLVLPASEATK 1453

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3082	37.9	1436	1	WC11_BOVIN
2	822.5	10.1	875	1	NETR_HUMAN
3	719.5	8.8	761	1	NETR_MOUSE
4	691.5	8.5	347	1	CD5L_HUMAN
5	664.5	8.2	352	1	CD5L_MOUSE
6	633.5	7.8	757	1	LQ4_MOUSE
7	629	7.7	754	1	LQ3_MOUSE
8	622.5	7.6	753	1	LQ3_HUMAN
9	618	7.6	756	1	LQ4_HUMAN
10	614.5	7.6	665	1	CD6_MOUSE
11	607.5	7.5	774	1	LQ2_HUMAN
12	592	7.3	668	1	CD6_HUMAN
13	520	6.4	532	1	SPER_STRPU
14	298	3.7	483	1	MRCO_MESAU
15	290.5	3.6	518	1	MTCO_MOUSE
16	290	3.6	458	1	MSRE_MOUSE
17	288.5	3.5	520	1	MRCO_HUMAN
18	288	3.5	451	1	MSRE_HUMAN
19	279.5	3.4	453	1	MSRE_BOVIN
20	279.5	3.4	454	1	MSRE_RABIT
21	264	3.2	494	1	CD5_MOUSE
22	258.5	3.2	2470	1	NTC2_MOUSE
23	258	3.2	5376	1	2AN_MOUSE
24	254.5	3.1	495	1	CD5_HUMAN
25	248.5	3.1	2471	1	NTC2_HUMAN
26	247.5	3.0	2471	1	NTC2_RAT
27	246.5	3.0	2531	1	NTC1_MOUSE
28	244.5	3.0	2871	1	FBN1_MOUSE
29	242.5	3.0	491	1	CD5_RAT
30	240.5	3.0	1238	1	JAG2_HUMAN
31	239.5	2.9	4544	1	LRP1_HUMAN
32	239	2.9	2531	1	NTC1_RAT
33	233	2.9	1700	1	BAR3_CHITE

RESULT 1

ID	WC11_BOVIN	STANDARD;	PRT;	1436 AA.
AC	P30205;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DE	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Antigen WC1.1 precursor.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID-9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Blood;			
RA	MEDLINE=93056489; PubMed=1431105;			
RA	Wijngaard P.B.J.; Metzelaar M.J.; Machugh N.D.; Morrison W.I.;			
RA	Clevers H.C.;			
RT	"Molecular characterization of the WC1 antigen expressed specifically on bovine CD4-CD8-gamma delta T lymphocytes.";			
RL	J. Immunol. 149:3273-3277(1992).			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- SIMILARITY: CONTAINS 11 SRCR DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL: X63723; CAA45255.1; -			
DR	PIR: S19913; S19913.			
DR	PIR: A46496; A46496.			
DR	InterPro: IPR001190; Srcr_receptor.			
DR	Pfam: PF00530; SRCR; 11.			
DR	PRINTS; PR00258; SPERACTRCPTR.			
DR	SMART: SM00202; SR: 11.			
DR	PROSITE; PS00420; SRCR_1; 4.			
DR	PROSITE; PS0287; SRCR_2; 11.			
KW	Antigen; Repeat; Glycoprotein; Signal.			
FT	SIGNAL 1 25 POTENTIAL.			
FT	CHAIN 26 1436 ANTIGEN WC1.1.			
FT	DOMAIN 28 131 SRCR 1.			
FT	DOMAIN 134 234 SRCR 2.			
FT	DOMAIN 239 340 SRCR 3.			
FT	DOMAIN 376 476 SRCR 4.			
FT	DOMAIN 481 581 SRCR 5.			
FT	DOMAIN 586 686 SRCR 6.			
FT	DOMAIN 689 789 SRCR 7.			
FT	DOMAIN 794 895 SRCR 8.			
FT	DOMAIN 931 1031 SRCR 9.			
FT	DOMAIN 1036 1136 SRCR 10.			
FT	DOMAIN 1155 1255 SRCR 11.			

P35555 homo sapien
P46531 homo sapien
P07207 drosophila
P98133 bos taurus
P97607 rattus norv
Q9tv36 sus scrofa
P61555 mus musculu
P19238 bos taurus
P35556 homo sapien
P98157 gallus gall
Q61982 mus musculu
Q9qy5 mus musculu

ALIGNMENTS

FT CARBOHYD 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 244 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 256 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 424 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 499 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 531 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 717 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 799 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 897 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 979 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 999 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1054 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1086 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1173 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1393 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1436 AA; 154196 MW; 6C72E5FDC6BA088C CRC64;
Query Match 37.9%; Score 3082; DB 1: Length 1436;
Best Local Similarity 41.7%; Pred. No. 1.2e-208;
Matches 613; Conservative 208; Mismatches 495; Indels 154; Gaps 23;
QY 28 CILLNSCEPLISFNGTDLRLVNGDPCSGTVEYKFGQWGTVCDDGWNNTASTVYCK 87
DB 13 CVLLG-----TWGQALELRKDGVRHCEGRVEYKHGEGWGTVDGYRWTLKDASVVC 67
QY 88 QLCGPFAM----FRGQAVTRHGIWLDVSCYGNESALWECQH---REWSHCYHGE 141
DB 68 QLCGCAAGIIFPGGAYFGPL--GPIWLLYTSCEGTSTVSDCEHSNIDKRYNDGYNHR 124
QY 142 DVGVCYGEANLGLRLVDGNNCSGRVEYKFOERWCTICDDGWNLTAAVVCROLGCPSS 201
DB 125 DAGVCSG----FVRLAGDGCPSGRVEYHSGEAMIPVSDGNTLATAOILCAELGCGKA 180
QY 202 FISSGVNVPAYLRPIWLLDILCOGNELALNCRHGRGNWHDCHSHNEDVTLTCYDSSDLE 261
DB 181 VSVLGHLPRESSAQWAEERCEGEPELWVCPVPCPGGTCHHSGSAQVVCYSEVR 240
QY 262 LRLVGTNRGMRVELKIOGRNGTVCHHKNWNAADYVCKQLGCGTALHFAGLPHLOSQ 321
DB 241 L-MTNGSSCOCEGOVEMNTSGORALCASHWSLANANVICROLGCGVAISTPGPHLVEEG 299
QY 322 DVVWLDGVSCNESPLDCRISGTVNFCDLHNDVSVICSDGADLELRADGNNCSGR 381
DB 300 DQILTARFHCSEFLWSCPVTALGGPDCSHGNTASVICSN----- 342
QY 382 VEVRIHQWTTCDONKWEQALVVCQLGCPFSVFGSRRKAPSNEARDIWNISICTGN 441
DB 343 -QIVLPQ-----CNDV-----SQTGSAAS----- 363
QY 442 ESALWDCYDGAKRKTCPRRSRSDAGVICSADLRLVGAHSPCYGRLEVKYQGEWTV 501
DB 364 -----EDAPYCSDR--QLRLDGGGPCAGRVEILDQGSWGTIC 401
QY 502 HDRWSTRNAVVCYKQLGCKPMPHFGMTYFKEASGPIWLLDVSICGNESNIWDCEHSGW 561
DB 402 DGGWLLDARVVCROLGCGCEALNATGSAHFGAGSGPIWLLNLCNCTKESHVWRCPGRWG 461
QY 562 KNCVHREDVYVTCSDATWGLRLVGGSNRCSGRLEVEYFGGRWGTVCDDGWSKAAVVC 621
DB 462 QHNCRHQDAGVICS--EFLALRWSEDOQCAGWLEVFYNGWTSVCRNPMEDITVSTIC 519
QY 622 SOLDCPSSIIIGLGNAST---GYCKIWLDDVSCDGEDSLWSCNMGWNNDCSHSD 677
DB 520 ROLGCGDS-----GTLNNSVALREGRPOWDRICQCKTDSLWQCPSPDPWNNNSCPREE 575
QY 678 VGVICSADSMELRLVGGSSRCAGKVEVNVQAVGTILCANGMNIATVYVQLCGCSAI 737
DB 576 AYIWCADS--RQIRLVGGGRCGRVEILDQGSWGTICDDRWLLDARVVCQLGCGEAL 633
QY 738 RVSREPHTERTLHILMNSGCGTGGASLWDCIRWEWKQTACHLNNEASLICSARQRL 797

DB 634 DATVSSFFCTGSGPIWLDEVNCRGEEQVWRCPSGWRQHNCNHQEDAGVICSGP--VRL 691
QY 798 VCADMPCSGRVEYKHADTWRSVCDSDFSLHAANVLCRELNCGDALSLSDVDFHGGNGLT 857
DB 692 AGDGPDCSGRVEYHSGEAWTPVSDGNFTLPTAQVICAELGCGKASVVLGHMPFRESGDV 751
QY 858 WAEKQCSESETHALCPIVQHPEDTCIHSREVGVVCSRYTDVRLV--NGKSQCDGQVEIN 916
DB 752 WAEKFRCDGGEPELWSCPRVPCPGCTCLHSGAAQVVCVYTEVQLMKNKTSQCEGVEMK 811
QY 917 VLGHWSGLDTHWDPEDARVLCRQLSCGTALTSTTGKYGIGERSVVRWGHFRHCLGNESLL 976
DB 812 ISGRWRALCASHWSLANANVVCRLGCGVAISTPRGPHLVEGGDQISTAQFHCSEAESL 871
QY 977 DNQOMTVLGAAPPCHHNTVSVICTGSLTQPLPCLANVSDPYLSAVPEGSALICLEDKRL 1036
DB 872 WSCPVTALGGPDCSHGNTASVCSGNHTQVLPQCNDFLSQPAGSRAESESPPYCSDSRL 931
QY 1037 RLVDGDSRCAGRVEYHIDGFWGTICDDGWDLSDAHVVCOKLGGVAFNATVSAHFGEGSG 1096
DB 932 RLVDGGGCGGRVEILDQGSWGTICDDDDWLLDARVVCRLGCGEALNATGSAHFGAGSG 991
QY 1097 PIWLDLNCCTGTESHLMQCPSPRGWQHDCRHKEDAGVICSEFTALRLYSETETESAGRL 1156
DB 992 PIWLDLNCCTGTESHLMQCPSPRGWQHDCRHKEDAGVICSEFTALRMVS--EDQCCAGWL 1049
QY 1157 EYFNGTWSVGRNRITTAIAGIVCROLCGCGGVVSLAPLSKGTSGFMWVDDIOCPKTH 1216
DB 1050 EYFNGTWSVGRNPMEDITVSVICROLGCGDSGLNTSVGLREGRSPRWVDDIOCRWD. 1109
QY 1217 ISIWOLCLAPWERRISSPAETWITCE-----DRIRVRGGDTESGRVE 1260
DB 1110 TSLWQCPSPGWKYSKSPKEAYISCEGRPKSCPTAACTDREKRLRGSGSECSGRVE 1169
QY 1261 IWHAGSWGTVCDSDWDLAEAEVVCQOLGCGSALALRDASFQGTGTIWLDMRCKNGNS 1320
DB 1170 VWHNGSWGTVCDSDWDLAEAEVVCQOLGCGSALAEVRSAAFGPGNGSIWLDEYVQCGG 1229
QY 1321 FLWDCHAKPWGSDCGHCKEDAGVRCSG-----QSLKSLNASSGHLLALIL 1364
DB 1230 SLWDCAEPWGGSDCKHEEDAGVRCSGVTTPTTAGTRTTSNLSLPGIFSPGLCLIL 1289
QY 1365 SSIFGLLLVLFLFTWCRVQKQKHLPLRVSTRRGSLEENLFHEMETCLKREDPHGTR 1424
DB 1290 GSLLFLVLVLTQLRW--RAER-----RALSSVEDALAEAYEBLYLTQKEGLGSP 1342
QY 1425 TS-DOTPNHGCEDASTSLGLVLPASEATK 1453
DB 1343 DQMTDVPDENYDDAEVVPVPTSPSQGNE 1372
RESULT 2
NETR_HUMAN
ID NETR_HUMAN STANDARD; PRT; 875 AA.
AC P56730; Q9UP16;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neurotysin precursor (EC 3.4.21.-) (Motopsin) (Leydin).
GN PRS12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98201705; PubMed=9540828;
RA Proba K., Gschwend T.P., Sonderegger P.;
RT "Cloning and sequencing of the cDNA encoding human neurotysin.";
RL Biochim. Biophys. Acta 1396:143-147(1998).
RN [2]

RC TISSUE=Brain;
 RA MEDLINE=97401523; PubMed=9245503;
 RX Gschwend T.P., Krueger S.R., Kozlov S.V., Wolfer D.P., Sonderegger P.,
 RT "Neurotrophin, a novel multidomain serine protease expressed in the
 RT nervous system.",
 RL Mol. Cell. Neurosci. 9:207-219(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98008848; PubMed=9344839;
 RA Yamamura Y., Yamashiro K., Tsuruoka N., Nakazato H., Tsujimura A.,
 RT Yamaguchi N.,
 RT "Molecular cloning of a novel brain-specific serine protease with a
 RT kringle-like structure and three scavenger receptor cysteine-rich
 RT motifs.",
 RL Biochem. Biophys. Res. Commun. 239:386-392(1997).
 CC -!- FUNCTION: PLAYS A ROLE IN NEURONAL PLASTICITY AND THE PROTEOLYTIC
 CC ACTION MAY SUBSERVE STRUCTURAL REORGANIZATIONS ASSOCIATED WITH
 CC LEARNING AND MEMORY OPERATIONS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: MOST ABUNDANT IN CEREBRAL CORTEX, HIPPOCAMPUS
 CC AND AMYGDALA.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 3 SRCR DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Y13192; CAA73646.1; .
 DR EMBL; D89871; BAA23986.1; .
 DR HSSP; P00763; 1DPO.
 DR MEROPS; S01.237; .
 DR MGD; MGI:1100881; Prss12.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR InterPro; IPR001190; Srcr_receptor.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR Pfam; PF00530; SRCR; 3.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE; FALSE_NEG.
 DR PRINTS; PR00258; SPERACTRCPT.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00202; SR; 3.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; FALSE_NEG.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS00420; SRCR_1; 3.
 DR PROSITE; PS0287; SRCR_2; 3.
 DR PROSITE; PS0240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 KW Hydrolase; Serine protease; Glycoprotein; Kringle; Repeat; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 761
 FT DOMAIN 85 157
 FT DOMAIN 166 267
 FT DOMAIN 273 373
 FT DOMAIN 386 487
 FT DOMAIN 505 761
 FT DOMAIN 505 516
 FT ACT_SITE 516 517
 FT ACT_SITE 562 562
 FT ACT_SITE 612 612
 FT ACT_SITE 711 711
 FT DISULFID 505 636

FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 521 521 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 569 569 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 761 AA; 84118 MW; DF507B03712164E6 CRC64;
 Query Match 8.8%; Score 719.5; DB 1; Length 761;
 Best Local Similarity 31.8%; Pred. No. 4.8e-43;
 Matches 191; Conservative 80; Mismatches 233; Indels 97; Gaps 19;
 QY 451 DCAKRTKFRSDAGVI-----CSDKADLD-LRLVGAHSPCYGRLEVKYQGEWGTVCVD 503
 DB 134 DSGRPWCIFYRNAQCKVDWGYDCGCPALPVIRLVGNSGHEGRVELYHAGWGTCID 193
 QY 504 RSTNAAVVCKQCGKPMHVFMTYFKEASGPITLWDDVSCIGNESNTWDCHEHGWKH 563
 DB 194 QMNDADADVICRLGLSIGIAKAWHAHFGSGSPILLDEVRCITNELSTEQPKSWGSH 253
 QY 564 NCVRHEDYIVTCSGDTWGLRLVGGSNCRSGRLEVYFQGRWGTVCDDGWSKAAAVVCSQ 623
 DB 254 NCGKHEDAGVCPITDGVIRLAGGKSTHEGRLEVYFGQWGTVCDDGTENTYVAC-- 311
 QY 624 LDCPSIIGMGLGNAST-----GYGK-IWLDVSCDGDSDLSWSCRNSGWNDCSHSD 677
 DB 312 -----RLLGFKYKQSSVNHFDGSRNPITWDDVSCGREVSFIQCSRRQWRHDCSHRD 366
 QY 678 VGVICSDASD-----NELRVGGSSRCAGKVEVNVGAVGILCANGWGMIAEVCVRQ 730
 DB 367 VGLTCYPDSGHRSLSPGFPPIRLVDGENKKEGRVEVFGWGTICDDGWTDKHAAVICRQ 426
 QY 731 LECGSAIRVSRPHPTERTLHILMSNGCTGGEASLWDCIRWEWKOTACHLNNEASLICS 790
 DB 427 LGYKPARARTWAYGEGKPIHMDNVKCTGNEKALADCVKODIGHNCRHSEDAGVICD 486
 QY 791 -----AHRQRLVGADMPGCSG-----RVEYKHADTWRS 818
 DB 487 YLEKKASSGNKEMLSGCGRLHRRQKRIITGGNSLRGAWPQASRLRSAGD-GR 545
 QY 819 VC-----DSDFSLHAANVLCRELNCDAISLSVGDH-----FGKNG- -LTWAEKQ 863
 DB 546 LCGATLLSSCWVITAAHCFKRYGNNSRYAVRVGDYHTLVPEEFQEIQQVQIVLHNR 605
 QY 864 CSGSETHALCPVQHPEDTCIH-SREVGVCVSRVTDVRLVNGKSCQDCQVEINVLGHW 922
 DB 606 PDRSDYDIALVRL-QGPGQCARLSTHVLPACLPLWRERPQKTASNC-----HITG-WG 657
 QY 923 SLCDTHWDPEDARVLCRLSCGTALSTGTGKYIGERSVWVWGHFRHCLGN-----ESL 979
 DB 658 -----DTGRAYSRITLQ-QAAVPLLPKFCRKERYKGLFTGRMLCAGNLOEDNRVDC 707
 QY 980 Q 980
 DB 708 Q 708
 RESULT 4
 CD5L_HUMAN
 ID CD5L_HUMAN STANDARD; PRT; 347 AA.
 AC 043866;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE CD5 antigen-like precursor (Sp-alpha) (CT-2) (Igm-associated peptide).
 GN CD5L.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., SEQUENCE OF N-TERMINUS, AND TISSUE SPECIFICITY.
 RC TISSUE=Splice;
 RA MEDLINE=9719777; PubMed=9045627;
 RA Gebe J.A., Kiener P.A., Ring H.Z., Li X., Francke U., Aruffo A.;
 RT "Molecular cloning, mapping to human chromosome 1 q21-q23, and cell

RESULT, T 6

Db 262 CQVAVPGRGKLRPACPGHIAVYVSVAGPHFRQKPKTKESHAEELKVLRLSGAQQV 321
QY 272 MGRVELKIQGRGWTVCHHKWNNAADVVCKQLGCGTALHFAGLPHLQSGDVVWLDGVSC 331
Db 322 EGRVFLNRQWGTVCRRNLISAVVCKQLGFGSREALFGAQLGCGLPILHSEVRC 381
-QY 332 SGNESFLWDC-RHSGTVNFCLHONDVSVICS---DGADLELRADSNCSGRVEVRH 387
Db 382 RGYERTLGDCLALGSON-GCQHANDAARVRCNIPDMGFQNKVRLAGGRNSEEYVVEQVE 440
QY 388 ----EQWWTICDQWKNNEQALVCKQLGCPFSVFGSRA---KPSNEARDIWNISCTG 440
Db 441 VNGGPRNGTVCSDHMGTEAMVTQRLGLGFANFALDWTWQGTPEAKVMSGVRCSG 500
QY 441 NESALWDCITVDGKAKRTCFR---RSDAGVICSDKA 472
Db 501 TEMALQOCQRHGPVH---CSHGPGFSGAGVACMNSA 533

RESULT 7
LOL3_MOUSE STANDARD; PRT; 754 AA.
AC Q22175; Q9JJ39;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lysyl oxidase homolog 3 precursor (EC 1.4.3.-) (Lysyl oxidase-like protein 3) (Lysyl oxidase related protein 2).
DE LOXL3 OR LOXL2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J, and 129/SvJ; TISSUE=Muscle;
RX MEDLINE=99126643; PubMed=9927484;
RA Jang W., Hua A., Spilson S.V., Miller W., Roe B.A., Meisler M.H.;
RT "Comparative sequence of human and mouse BAC clones from the mnd2 region of chromosome 2p13.";
RL Genome Res. 9:53-61(1999).
CC -1- COFACTOR: COPPER AND LYSINE TYROSYLQUINONE (LTQ) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Extracellular (Potential).
CC -1- PTM: THE LYSINE-TYROSYLQUINONE IS GENERATED BY THE MODIFICATION OF A TYROSINE RESIDUE CROSSLINKED TO THE EPSILON-AMINO GROUP OF A LYSINE (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 4 SRCR DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE LYSYL OXIDASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL: AF053368; AAC83205.1; -
CC EMBL: AF084363; AAC95338.1; -
CC MGD: MGI:1337004; Loxl3.
CC InterPro: IPR001695; Lysyl_oxidase.
CC InterPro: IPR001190; Srcr_receptor.
CC Pfam: PF00530; SRCR; 4.
CC Pfam: PF01186; Lysyl_oxidase; 1.
CC PRINTS: PR00074; Lysyl_oxidase.
CC PRINTS: PR00258; SPERACTRPT.
CC ProDom: PD013887; Lysyl_oxidase; 1.
CC SMART: SM00202; SR; 4.
CC PROSITE: PS00926; LYSYL_OXIDASE; 1.
CC PROSITE: PS00420; SRCR_1; 1.
CC PROSITE: PS00287; SRCR_2; 4.
KW Oxidoreductase; Copper; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 26 POTENTIAL.

FT CHAIN 27 754 LYSYL OXIDASE HOMOLOG 3.
FT DOMAIN 45 146 SRCR 1.
FT DOMAIN 170 283 SRCR 2.
FT DOMAIN 308 408 SRCR 3.
FT DOMAIN 418 526 SRCR 4.
FT DOMAIN 530 733 LYSYL-OXIDASE LIKE.
FT METAL 608 608 COPPER (POTENTIAL).
FT METAL 610 610 COPPER (POTENTIAL).
FT METAL 612 612 COPPER (POTENTIAL).
FT MOD_RES 635 635 CROSS-LINKED TO TYROSYLQUINONE (BY SIMILARITY).
FT MOD_RES 671 671 TYROSYLQUINONE (BY SIMILARITY).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 391 391 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 626 626 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 754 AA; 83681 MW; F8758400943F4FF2 CRC64;
Query Match 7.7%; Score 629; DB 1; Length 754;
Best Local Similarity 31.0%; Pred. No. 1.1e-36;
Matches 155; Conservative 72; Mismatches 183; Indels 90; Gaps 16;
QY 56 FCSGTVEYKFOGQRTVCDGNTTASTVCKQLGCPFSFA-----MFRFOAVTRHGKI 110
Db 54 FYEGRVBIQIRAGEWGTICDDDFTLQAAHVLCRELG--FTEATGWTHSAKYPGT---GRI 108
QY 111 WLDVSCYGNESALWECQHWGSHNCHYGEDVGVNC-----YGEANL----- 153
Db 109 WLDNLSRGTEGSGYTECASRGWNSDCTHDEDAGVICKDRLPGFSDSNVIEVHQLOVE 168
QY 154 -----GLRLVDGNNSCSGRVEYKFOERWGTICDDGNLNTAAVVCQLGCPSS-- 201
Db 169 EVRLRPVAVGRRPLP---VTEGLVEVRLPGWQVQCDKWSAHNSHVVCMLGFPGEKR 225
QY 202 -----FISGGVYNPAVLRIWLDLILCOGNELALMNCRRHGWGND---CSHNEDVT 251
Db 226 VNMAFYRLMAQKKQHS-----FGLHSVACVTEAHLSCLSLEFYRANDTTTRCSGPNPAV 279
QY 252 LTCY-----DSSDLELRVLGVTNCRMGVRLKICGRWCTVCHHKWN 293
Db 280 VSCVLGPLYAFYFTGOKKQKHPOGEARVLRUKGAHOGEGRVEVLKAGTWGTVCDKRWDL 339
QY 294 AAADVCKQLCGTALHFAGLPHLQSGSDVVWLDGVSCSGNESFLWDCRHSCTVNFDCILH 353
Db 340 QAASVVCPELGFGTAREALSARGMQCGNGAIHLSEVRCSGOEPLSWRCPSKNITAECDCH 399
QY 354 QNDVSVICS---DGADLELRADSNCSGRVEVRI-----HEQWWTICDQWKNNEQALV 406
Db 400 SDQAGVRCNLPTGYETKIRLSGGRSRYEGRVEVQIGIPGHLRWGLICGDDMGTLKAVYA 459
QY 407 CKQLGCPFSVFGSRA---KPSNEARDIWNISCTGNESALWDCITVDGKAKRTCFR---R 461
Db 460 CRQLGLGVANHLGLETWYDWSGNTVEVVMVSGVRCITGSELSNQCAHF--SSHITCKKTGR 518
QY 462 SDAGVICSKADLRLRLVGA 481
Db 519 FTAGVICSETAS-DLLLSHA 537
RESULT 8
LOL3_HUMAN STANDARD; PRT; 753 AA.
AC P58215; Q96RS1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lysyl oxidase homolog 3 precursor (EC 1.4.3.-) (Lysyl oxidase-like protein 3).
DE LOXL3 OR LOXL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21181571; PubMed=11284725;
RA Maki J.M., Kivirikko K.I.;
RT "Cloning and characterization of a fourth human lysyl oxidase
isoenzyme";
RL Biochem. J. 355:381-387(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21280915; PubMed=11386757;
RA Jourdan-Le Saux C., Tomshe A., Ufalusi A., Jia L., Csizsar K.;
RT "Central nervous system, uterus, heart, and leukocyte expression of
the LOXL3 gene, encoding a novel lysyl oxidase-like protein.";
RL Genomics 74:211-218(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21233589; PubMed=11334717;
RA Huang Y., Dai J., Tang R., Zhao W., Zhou Z., Wang W., Ying K., Xie Y.,
Mao Y.;
RT "Cloning and characterization of a human lysyl oxidase-like 3 gene
(hLOXL3).";
RL Matrix Biol. 20:153-157(2001).
CC -1- COPACTOR: COPPER AND LYSINE TYROSYLQUINONE (LTO) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Extracellular (Potential).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY TISSUES, THE HIGHEST LEVELS
AMONG THE TISSUES STUDIED BEING IN THE PLACENTA, HEART,
OVARY, TESTIS, SMALL INTESTINE AND SPLEEN.
CC -1- PTM: THE LYSINE-TYROSYLQUINONE IS GENERATED BY THE MODIFICATION OF
A TYROSINE RESIDUE CROSSLINKED TO THE EPSILON-AMINO GROUP OF A
LYSINE (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 4 SRCR DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE LYSYL OXIDASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL: AF282619; AAK51671.1; -;
CC EMBL: AF311313; AAK63205.1; -;
CC EMBL: AF284815; AAK91134.1; -;
CC GenBank: HGNC:13869; LOXL3.
CC InterPro: IPR001695; Lysyl_oxidase.
CC InterPro: IPR001190; Srcr_receptor.
CC Pfam: PF00530; SRCR; 4.
CC Pfam: PF01186; Lysyl_oxidase; 1.
CC PRINTS: PR00074; LYSILOXIDASE.
CC PRINTS: PR00258; SPERACTRCPTR.
CC ProDom: PD013887; Lysyl_oxidase; 1.
CC SMART: SM00202; SR; 4.
CC PROSITE: PS00926; LYSYL_OXIDASE; 1.
CC PROSITE: PS00420; SRCR_1; 1.
CC PROSITE: PS0287; SRCR_2; 4.
CC Oxidoreductase; Copper; Glycoprotein; Repeat; Signal.
KW SIGNAL 1 25
FT CHAIN 26 753
FT DOMAIN 44 145
FT DOMAIN 169 282
FT DOMAIN 307 407
FT DOMAIN 417 525
FT DOMAIN 529 732
FT METAL 607 607
FT METAL 609 609
FT METAL 611 611
FT MOD_RES 634 634
FT MOD_RES 670 670
FT CARBOHYD 111 111
FT CARBOHYD 266 266

FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 159 159 E -> K (IN REF. 3).
SQ SEQUENCE 753 AA; 83166 MW; 582C46DA25E05A69 CRC64;

Query Match 7.68; Score 622.5; DB 1; Length 753;
Best Local Similarity 31.28; Pred. No. 3.1e-36;
Matches 167; Conservative 71; Mismatches 202; Indels 95; Gaps 19;

QY 30 LLLNSCFLLISFNFGT-----DLELRVN-GDGPCSGTVEVKFGQWGTVD 74
DB 14 LLL--CLLCSSCLGSPSTGPEKAGSQGLRFLRAGPRKPYEGRVIQIRAGETICD 71
QY 75 DGNWNTASTVVKVQKLGCPFSFA-----MRFQGVATRIGKIWLDDVSCYNESALWECQH 129
DB 72 DDFTLQAAHILCRELG--FTEATGWTHTAKYGPGT--GRIWLDNLSCGTSEQSTECAS 126
QY 130 REMGSHNCHGEDVGVNC-----YGEANL-----GLRLVDGNN 162
DB 127 RGWNSDCDTHEDAGVICKDORLPGFSDSNVIEVEHHLOVEVRIRPAVGWGRRLP 183
QY 163 SCGRVEYKFORWCTICDDGNLNTAAVVCRLQPCSSFISSGVVNSPAVL 219
DB 184 VTEGLVEVRLPDGWSQVCDKGSASHNVCGMLGFPSEKRVNAAFYRLAQROHSEGL 243
QY 220 DDLCQGNELALWNCRHGRGNHD---CSHNEDEVLTLC-----YDSS 258
DB 244 HGVACVGEAHLSLGLEFYRANDTARCPGGPVPVPGPVYAASSGQKQKQSKPG 303
QY 259 DLELRVGVNRCMGRVELKIQGRWGTVCHHKWNAADVVCOLGCGCTALHAGLPHLQ 318
DB 304 EARVRLKGAHPGEGRVEVLKASTWGTCDRKWDHAAASVVCRELGFSGARALSGRMG 363
QY 319 GSDVYVWLDVSCSGNESFLWDCRHSGTVNFDCLHQNVDVICS---DGADLELRADGS 375
DB 364 QGMGAILHSEVRCSGQLSLNCKPHKNITAECDSSHQDAGVRCNLPTGTAEIRLSGGR 423
QY 376 NNCGRVEVRIHE-----QWTTCDQNKNEQALVCKQLGCPFSVFGSRA--KFSNEAR 429
DB 424 SQHEGRVEYQIGPGPLRMGLGICDDWGTLEAMVACRQLGLGYANHLQETWYDWSGNIT 483
QY 430 DIWINSISCTGNESALWDCTYDGKAKRTCFR---RSDAGVICSKADLDRLVGA 481
DB 484 EVVMGVRCTGTGTELSLDQCAHNG--THICTKRTGTRTAGVICSETAS-DLLHSA 536

RESULT 9
LOLA_HUMAN STANDARD; PRT; 756 AA.
ID LOLA_HUMAN
AC Q96JB6; Q96PC0; Q96DY1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lysyl oxidase homolog 4 precursor (EC 1.4.3.-) (Lysyl oxidase-like
protein 4) (Lysyl oxidase related protein C).
GN LOXL4 OR LOXC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Testis;
RX MEDLINE=21316447; PubMed=11292829;
RA Ito H., Akiyama H., Iguchi H., Iyama K., Miyamoto M., Ohsawa K.,
Nakamura T.;
RT "Molecular cloning and biological activity of a novel lysyl oxidase-
related gene expressed in cartilage";
RL J. Biol. Chem. 276:24023-24029(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21550107; PubMed=11691589;

RA Maki J.M., Tikkanen H., Kivirikko K.I.;
 RT "Cloning and characterization of a fifth human lysyl oxidase
 RT isoenzyme: the third member of the lysyl oxidase-related subfamily
 RT with four scavenger receptor cysteine-rich domains.";
 RL Matrix Biol. 20:493-496(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Asuncion L.P., Fogelgren B., Fong K.S.K., Fong S.F.T., Kim Y.,
 RA Csizsar K.;
 RT "A novel human lysyl oxidase-like gene (LOXL4) on chromosome 10q24 has
 RT an altered SRCR domain and is down-regulated by H-ras.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Eye;
 RA Strausberg R.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: May modulate the formation of a collagenous
 CC extracellular matrix.
 CC -I- COFACTOR: COPPER AND LYSINE TYROSYLQUINONE (LTO) (BY SIMILARITY).
 CC -I- SUBCELLULAR LOCATION: Extracellular (Potential).
 CC -I- TISSUE SPECIFICITY: Expressed in many tissues, the highest levels
 CC among the tissues studied being in the skeletal muscle, testis and
 CC pancreas. Expressed in cartilage.
 CC -I- PTM: THE LYSINE-TYROSYLQUINONE IS GENERATED BY THE MODIFICATION OF
 CC A TYROSYL RESIDUE CROSSLINKED TO THE EPSILON-AMINO GROUP OF A
 CC LYSINE (BY SIMILARITY).
 CC -I- SIMILARITY: CONTAINS 4 SRCR DOMAINS.
 CC -I- SIMILARITY: BELONGS TO THE LYSYL OXIDASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF38441; AAK71934.1; -
 CC EMBL: AY036093; AAK64186.1; -
 CC EMBL: AF395336; AAL27543.1; -
 CC EMBL: BC013153; AAH13153.1; -
 CC Genew: HGNC:17171; LOXL4.
 CC InterPro: IPR001695; Lysyl_oxidase.
 CC InterPro: IPR001190; Srcr_receptor.
 CC Pfam: PF01186; Lysyl_oxidase; 1.
 CC Pfam: PF00530; SRCR; 4.
 CC ProDom: PD013887; Lysyl_oxidase; 1.
 CC PROSITE: PS00926; LYSYL_OXIDASE; FALSE_NEG.
 CC PROSITE: PS00420; SRCR.1; 1.
 CC PROSITE: PS50287; SRCR.2; 4.
 CC Oxidoreductase; Copper; Glycoprotein; Repeat; Signal.
 KW SIGNAL
 FT 1 24 POTENTIAL.
 FT CHAIN 25 756 LYSYL OXIDASE HOMOLOG 4.
 FT DOMAIN 32 133 SRCR 1.
 FT DOMAIN 159 287 SRCR 2.
 FT DOMAIN 311 411 SRCR 3.
 FT DOMAIN 421 529 SRCR 4.
 FT DOMAIN 533 736 LYSYL-OXIDASE LIKE.
 FT METAL 611 611 COPPER (POTENTIAL).
 FT METAL 613 613 COPPER (POTENTIAL).
 FT METAL 615 615 COPPER (POTENTIAL).
 FT MOD_RES 638 638 CROSS-LINKED TO TYROSYLQUINONE (BY
 FT SIMILARITY).
 FT MOD_RES 674 674 TYROSYLQUINONE (BY SIMILARITY).
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 3 3 W -> R (IN REF. 4).
 FT CONFLICT 101 101 R -> Q (IN REF. 4).
 FT CONFLICT 405 405 D -> A (IN REF. 4).
 FT CONFLICT 493 493 S -> G (IN REF. 3).
 FT CONFLICT 539 539 A -> T (IN REF. 3).

FT CONFLICT 542 542 V -> A (IN REF. 3).
 FT CONFLICT 703 703 Y -> H (IN REF. 3).
 SQ SEQUENCE 756 AA; 84483 MW; 13051ACADB922BBC CRC64;
 Query Match 7.6%; Score 618; DB 1; Length 756;
 Best Local Similarity 31.2%; Pred. No. 6.4e-36;
 Matches 159; Conservative 68; Mismatches 191; Indels 92; Gaps 16;
 QY 47 ELRLVNDG-PCSGTVEVKFGQGTGTCDDGNTTASTVVKOLGCPFSFAM-----FRF 100
 Db 31 KLRLLGPESPEGRLEVLHGQGTGTCDDNFALQEAQVACRQLG--FEALTWHAHSAKY 88
 QY 101 GOAVTRGKIWLDDVSCYGNESALWECQHREWSHNCYHGEDGVNVCYGEANGLRLVDG 160
 Db 89 GQG---EGPIWLDNVRVCVGTSSLDQCGSNMGVSDCSHSDYGVICHPRRHRYLSETV 145
 QY 161 NNS-----CSGRVEVFQERWGTICDDGNLNTAAVVCV 194
 Db 146 SNALGPQGRLEEVRLKPIILASAKQHSPTVEGAVKRYEGHWRQVCDQGTMMNSRYVCG 205
 QY 195 QLGPSSF-ISS-----GVNSPAVLRLPIWLDLILCOGNELALWNCR--- 235
 Db 206 MLGFPSEVPVDVSHYRKYRWDLKMRDPKSLKSLTNKNSFWIHOVTCVLCGTEPHMANCOVQV 265
 QY 236 --HRGNHNDCHSHNEDVTLTCYD-----SSDLELRVLGGTNRMGV 276
 Db 266 APARKLRPACPGMHAVVSCVAGHPRPKTPQRKGSWAEPRVRLRSGAQVGEGRVE 325
 QY 277 LKIQGRWGTVCHHKWNNAAADVCKQLCGGTALHFAGLPHLOSQSDVVWLDGVSCSNGES 336
 Db 326 VLNRQWGTGCDHRNLIISAVVCRQLGFGSAREALFGARLQGLGPIHLSEVRCRGER 385
 QY 337 FLWDC-RHSQTVNFCLQNDVSVIC---SDGADLELRADGNSNCSGRVEVRTH---E 388
 Db 386 TLDCEPALESQN-GCOHENDAAVRNVNPNWGFONQVRLAGRIPEGLLEQVVEVNGVP 444
 QY 389 QWMTICDQNNKNEQALVYCKQLGCPFSVFGSRRRA---KPSNEARDIWNISCTGNSAL 445
 Db 445 RWGSCVSENGLTEAMVACVRLGLGLGFAIHAYKETWMSGTPRAQEVVNVMSGVRCSGTTEL 504
 QY 446 WDCYDQKAKRTCFER---RSDAGVICSOKA 472
 Db 505 QQCQRHGPVH--CSHGGRFLAGVSCMDSA 532
 RESULT 10
 CD6_MOUSE
 ID CD6_MOUSE STANDARD; PRT; 665 AA.
 AC Q61003; Q61004; Q60679;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE T-cell differentiation antigen CD6 precursor.
 GN CD6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RP TISSUE=Thymocytes;
 RX MEDLINE=96062291; PubMed=7594475;
 RA Robinson W.H., Prohaska S.S., Santoro J.C., Robinson H.L.,
 RA Parnes J.R.;
 RT "Identification of a mouse protein homologous to the human CD6 T cell
 RT surface protein and sequence of the corresponding cDNA.";
 RL J. Immunol. 155:4739-4748(1995).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RX MEDLINE=95174761; PubMed=7870060;
 RA Whitney G., Bowen M., Neubauer M., Aruffo A.;
 RT "Cloning and characterization of murine CD6.";
 RL Mol. Immunol. 32:89-92(1995).


```

or send an email to license@isb-sib.ch).
-----
CC EMBL; X60992; CAA43306.1; -
CC EMBL; U66142; AAC51161.1; -
DR DR EMBL; U66143; AAC51162.1; -
DR DR EMBL; U66144; AAC51163.1; -
DR DR EMBL; U66145; AAC51164.1; -
DR DR EMBL; U66146; AAC51165.1; -
DR DR EMBL; U66147; AAC51166.1; -
DR DR PIR; S26741; S26741.
DR DR MIM; H6720; -
DR DR MIM; H6720; -
DR DR InterPro; IPR001190; Srcr_receptor.
DR DR Pfam; PF00530; SRCR; 3.
DR DR PRINTS; PR00258; SPERACTRCPTR.
DR DR SMART; SM00202; SR; 3.
DR DR PROSITE; PS00420; SRCR_1; 1.
DR DR PROSITE; PS0287; SRCR_2; 3.
DR DR Cell adhesion; Glycoprotein; Transmembrane; Repeat; Signal;
KW Alternative splicing.
FT SIGNAL 1 17
FT CHAIN 18 668
FT DOMAIN 18 402
FT TRANSMEM 403 423
FT DOMAIN 424 668
FT DOMAIN 424 156
FT DOMAIN 161 260
FT DOMAIN 265 361
FT CARBOHYD 28 28
FT CARBOHYD 49 49
FT CARBOHYD 112 112
FT CARBOHYD 118 118
FT CARBOHYD 229 229
FT CARBOHYD 339 339
FT CARBOHYD 345 345
FT CARBOHYD 368 368
FT CARBOHYD 431 462
FT VARSPLIC 463 504
FT VARSPLIC 463 504
FT N -> D (IN ISOFORM CD6C AND ISOFORM
FT VMLPIQVAPPEDSDSDSDYEHYDFSAOPVPTFF
FT CD6E).
FT MISSING (IN ISOFORM CD6D AND ISOFORM
FT CD6E).
FT VMLPFI -> GPGAP (IN REF. 1).
FT MISSING (IN REF. 2; AAC51162).
FT SEQUENCE 668 AA; 71828 MW; EFB434608012C49E CRC64;

Query Match 7.3%; Score 592; DB 1; Length 668;
Best Local Similarity 28.7%; Pred. No. 3.7e-34;
Matches 154; Conservative 59; Mismatches 140; Indels 184; Gaps 15;

QY 900 VRLVNGKSCDQGVNVLGHWSGLCDTHWDPEDARVLCRLSCGTALSTGGKIYGRS 959
DB 45 VRLTNGSSSCGTVFVRLEASWEPAGCALWDSRAEAVCRALCGGAEASQ----- 96
QY 960 VRYWGRHFCL-----GNESLLDNCOMVLGAPCIHGNVTYSVICTGSLTOP 1006
DB 97 -----LAPPTPELPPPPAAGNTSVANA--FLAGAP-----ALLCSGAEWR- 135
QY 1007 LFPCLANVSDPYLSAYPEGSALICLEDKRLRLVDGSRACAGRVVEIYHDFGFWETICDDGWD 1066
DB 136 -----LCEVVEHACRSQDGRARRVTCARNRRLVLDGGACAGRVEMLEHGEWSVCDDTWD 191
QY 1067 LSDAHVVCQKLGCGVAFNATVSAHFEGSGPIWLDLNTGTGESHUWQCPSCRWGQDHC 1126
DB 192 LEDARVVCRLGCGWVAQALPGLHFTPGRPIHRDQVNGSABAYLWDCPGLP-GQHYCG 250
QY 1127 HKDAGVICSEFTALRLYSETETESACRLEVEYNGTWSVGRNITTAGIVCRLQGC 1186
DB 251 HKDAGVVCSE----- 261
QY 1187 GENGWVSLAPLSKTSFGFMWDDIQCPKTHISIWQCLSAFWERRISSPAEETWITCEDRI 1246
DB 262 -----HOSW----- 265

us-09-759-130b-381.rsp
-----
QY 1247 RVRGDTGTCGRVETWAGSHGTVCDSDWDLAEAEVVCQQLGCGSALAALRDASFGOGTG 1306
DB 266 RLTCGADRCCEQVEHFRGVWNTVCDSEWYPSKAVKVLCSLGCCTAVERPKGLPHSL-SG 324
QY 1307 TIWLDMMCKRGNESFLMDCHAKPWQSDCGKHEDAGVRCS-GQSLKSL----- 1353
DB 325 RMYYS---CNGEELTSCSWRFNNSLCSQSLAARVLCSSASRLNLSLSTPEVPASVOTV 381
QY 1354 -----NASSGHLALILSSI-FGLLLL--VLFLFTWCRVKQKHLPLRVS 1396
DB 382 TIESVTVKIENKESRELMLLIPSLIVIGILLGLSLFIATIL-LRIKGYALPVMVN 437

RESULT 13
ID SPER_STRPU STANDARD; PRT; 532 AA.
AC P16264; 1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Egg peptide speract receptor precursor.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OC NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 477-489.
RX MEDLINE=89184581; PubMed=2538832;
RA Dangott L.J., Jordan J.E., Bellet R.A., Garbers D.L.;
RT "Cloning of the mRNA for the protein that crosslinks to the egg
RT peptide speract."
RL Proc. Natl. Acad. Sci. U.S.A. 86:2128-2132(1989).
CC -!- FUNCTION: RECEPTOR FOR THE EGG PEPTIDE SPERACT.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: CONTAINS 4 SRCR DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J04518; AAA30078.1; -
DR PIR; A32751; A32751.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00530; SRCR; 4.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00202; SR; 4.
DR PROSITE; PS00420; SRCR_1; 4.
DR PROSITE; PS0287; SRCR_2; 4.
DR Receptor; Transmembrane; Glycoprotein; Signal; Repeat.
FT SIGNAL 1 30
FT CHAIN 31 532
FT DOMAIN 31 491
FT TRANSMEM 492 520
FT DOMAIN 521 532
FT DOMAIN 43 144
FT DOMAIN 153 257
FT DOMAIN 264 366
FT DOMAIN 382 485
FT CARBOHYD 78 78
FT CARBOHYD 115 115
FT CARBOHYD 459 459
FT SEQUENCE 532 AA; 57820 MW; 742533E095769CB8 CRC64;

Query Match 6.4%; Score 520; DB 1; Length 532;
Best Local Similarity 29.1%; Pred. No. 3.3e-29;
Matches 152; Conservative 66; Mismatches 208; Indels 96; Gaps 19;

```


QY 1 MMLPQNSWHIDFGRCCHQNLFSAVTTCILLNSFLISSFNGFDLE-----LR 49
Db 5 MMLQQYCW-----AACLVI---CIAISSVDDVGAQNTYGREAEGNIR 44
QY 50 LVNGDGPSCGTVEVKFOGQVCTVCDGWNNTASTVVCVKQLGCPSPFAMFR---FGQAVTR 106
Db 45 LIHGRTENESVEIYHATRHGGVCDWWHMHENANVTCKQLGFGPGARQFYRAYEGAHVT- 103
QY 107 HGKILWLDVSCYGNESALWECQHEMG--SHNCHGEDGVNVCY--GEANLGLRLVDGNN 163
Db 104 --TPWVKMKNCLGNETLEDYHPRYGRPWLCNAQWAGVECLPKDPEQSLRMLIGDVP 161
QY 164 CSGRVEVKFQERWGTICDDGNLNTAAVVCRLQGPCSPFSISSGVVNSPA-----VLRPI 217
Db 162 NEGTLFTWDGAMGSVCHTFDTPDGNVACRQM-----YSRGVKSITKIDGHFGFSTGPI 216
QY 218 WLDDILCOGNELALWNCRHGRWGN-----HDC--SHNEDVTLTCYDSSDLELRVGGT 268
Db 217 ILDAVDCGEGTAHITEC-----NMPVTPYOHACPYTHNMWDVGVCKPNVEGDRLMDGS 270
QY 269 NRCMGRVELKIQGRWGTVCCHKNNNAADVVCVKQLGCGTALHFAGLPHLOSGSDV----- 323
Db 271 GPHEGRVEIWHDDAWGTICDDGWDADANVVCROAGYRGVAKSGF---KGEFGFTWA 326
QY 324 -VWLDGVSCGNSFLWDCRHSGTVNFDCLHQNDVSVICDGADELRLAD-----G 374
Db 327 PIHTSFYMTGTVEVDRLICILRDGWTSHCYHVEDASVVCATDDDDTIEIEPKHTRVIRV 386
QY 375 SNNSCGRVEVRIHQWTTICDQNNKQALVVCVKQLGCPSPFSVGRRAKPSNEA----- 428
Db 387 MGQCGRVEVSLGNGWGRVCDPDMDSDEAKTVCYHAGYK---WGASRAAGSAEVSAPFDL 443
QY 429 -RDWISISCTGNES-ALMDCTYDGRKAKTCFRRSAGVIC 468
Db 444 EAPIIDGTCVGENETLSQCKMKVSADMTCT-ATGDVGVC 484

RESULT 14

MRCO_MESAU
ID MRCO_MESAU STANDARD; PRT; 483 AA.
AC QWUB9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Macrophage receptor MARCO (Macrophage receptor with collagenous structure).
GN MARCO.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99242649; PubMed=10224280;
RA Palecanda A., Paulauskis J., Al-Mutairi E., Imrich A., Qin G., Suzuki H., Kodama T., Tryggvason K., Koziel H., Kobzik L.;
RT "Role of the scavenger receptor MARCO in alveolar macrophage binding of unopsonized environmental particles";
RL J. Exp. Med. 189:1497-1506(1999).
CC -!- FUNCTION: Bind Gram-positive and Gram-negative bacteria.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein (Probable).
CC -!- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SRCR DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AF125191; AAD20360.1; -.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00530; SRCR; 1.
DR Pfam; PF01391; Collagen; 4.
DR PRINTS; PR00258; SPERACTRCPTR.
DR PRODOM; PD000007; Collagen; 1.
DR SMART; SM00202; SR; 1.
DR PROSITE; PS00420; SRCR_1; FALSE_NEG.
DR PROSITE; PS0287; SRCR_2; 1.
KW Collagen; Transmembrane; Receptor; Glycoprotein; Signal-anchor.
FT DOMAIN 1 48 CYTOPLASMIC (POTENTIAL).
FT FT 49 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT 69 (POTENTIAL).
FT FT 69 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 70 483 COLLAGEN-LIKE.
FT DOMAIN 148 383 SRCR.
FT DOMAIN 389 483 BY SIMILARITY.
FT DISULFID 412 472 BY SIMILARITY.
FT DISULFID 425 482 BY SIMILARITY.
FT DISULFID 452 462 BY SIMILARITY.
FT CARBOHYD 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 483 AA; 49621 MW; C38F18C46505FB1E CRC64;
Query Match 3.7%; Score 298; DB 1; Length 483;
Best Local Similarity 52.0%; Pred. No. 1.1e-13;
Matches 53; Conservative 13; Mismatches 30; Indels 6; Gaps 2;

QY 1034 KRLRLVDGDSRCAGREYTHDGFWTICDDGWDLSDAHVVCQKLCGVAFAFATVSAHGE 1093
Db 387 KEVIVGVTNR--GRAEIFYNNAGTICDDNNDNATVFCRMLG-----YSSGKGFTFG 440
QY 1094 GSGPIWLDLCTGESHLCWCPGSRGWHGHCRCHEKEDAGVIC 1135
Db 441 GSGNIWLDVNCQGTEDSLWNCRNKRWGSHNCNHNEDAGVEC 482

RESULT 15

MTCO_MOUSE
ID MTCO_MOUSE STANDARD; PRT; 518 AA.
AC Q60754;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Macrophage receptor MARCO (Macrophage receptor with collagenous structure).
GN MARCO.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95171455; PubMed=7867067;
RA Elomaa O., Kangas M., Sahlberg C., Tuukkanen J., Sormunen R., Liakka A., Thesleff I., Kraal G., Tryggvason K.;
RT "Cloning of a novel bacteria-binding receptor structurally related to scavenger receptors and expressed in a subset of macrophages";
RL Cell 80:603-609(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99265975; PubMed=10331948;
RA Kangas M., Brannstrom A., Elomaa O., Matsuda Y., Eddy R., Shows T.B., Tryggvason K.;
RT "Structure and chromosomal localization of the human and murine genes for the macrophage MARCO receptor";
RL Genomics 58:82-89(1999).
CC -!- FUNCTION: Bind Gram-positive and Gram-negative bacteria.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: Expressed in subpopulations of macrophages in the spleen and the medullary cord of lymph nodes.
CC -!- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SRCR DOMAIN.

```
CC ----- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U18424; AAA68638.1; .
DR EMBL; AF128423; AADS1136.1; .
DR EMBL; AF127927; AADS1136.1; JOINED.
DR EMBL; AF127928; AADS1136.1; JOINED.
DR EMBL; AF128169; AADS1136.1; JOINED.
DR EMBL; AF128170; AADS1136.1; JOINED.
DR EMBL; AF128171; AADS1136.1; JOINED.
DR EMBL; AF127601; AADS1136.1; JOINED.
DR EMBL; AF127602; AADS1136.1; JOINED.
DR EMBL; AF128419; AADS1136.1; JOINED.
DR EMBL; AF128420; AADS1136.1; JOINED.
DR EMBL; AF128421; AADS1136.1; JOINED.
DR EMBL; AF128422; AADS1136.1; JOINED.
DR MGD; MGI:130998; Marco.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR001190; Sctcr_receptor.
DR Pfam; PF00530; SRCR; 1.
DR Pfam; PF01391; Collagen; 5.
DR PRINTS; PR00258; SPERACTRCPTR.
DR ProDom; PD000007; Collagen; 1.
DR SMART; SM00202; SR; 1.
DR PROSITE; PS00420; SRCR_1; FALSE_NEG.
DR PROSITE; PS0287; SRCR_2; 1.
KW Collagen; Transmembrane; Receptor; Glycoprotein; Signal-anchor.
FT DOMAIN 1 48 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 49 69 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT (POTENTIAL).
FT DOMAIN 70 518 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 149 418 COLLAGEN-LIKE.
FT DOMAIN 423 518 SRCR.
FT DISULFID 446 507 BY SIMILARITY.
FT DISULFID 459 517 BY SIMILARITY.
FT DISULFID 487 497 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 518 AA; 52730 MW; B09E7601ECA23637 CRC64;

Query Match 3.6%; Score 290.5; DB 1; Length 518;
Best Local Similarity 49.5%; Pred. No. 4.2e-13;
Matches 51; Conservative 16; Mismatches 31; Indels 5; Gaps 2;

QY 1034 KRLRLVDGSRGAGVEIYHDFGFWGTCDDGWDLSDAHVVCKLGGGVAFNATVSAHFGE 1093
DB 421 QVRIRMGGTNR--GRAEVYNNWGTICDDNDNDATVFCRMIGYS---RGRALSSYGG 475

QY 1094 GSGPIWLDLNCGTGTFESHLWCQPSRGWGQDCHRHKEDAGVICS 1136
DB 476 GSGNIWLDNVNCGTENSILWDCSKNSGWNHNCVHNEDAGVECS 518
```

Search completed: May 12, 2003, 06:37:50
Job time : 28 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 06:36:46 ; Search time 24 Seconds
(without alignments)
1781.313 Million cell updates/sec

Title: US-09-759-130B-381
Perfect score: 8138
Sequence: 1 MMLPQNSHIDFGCCCHQ.....CEDASDTSLGLVPASEATK 1453

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2435	29.9	1785	4	US-09-341-587-3
2	1048.5	12.9	1290	1	US-08-470-350B-2
3	718.5	8.8	666	4	US-09-341-587-1
4	691.5	8.5	346	3	US-09-034-916-2
5	629	7.7	754	4	US-09-276-400-8
6	629	7.7	754	4	US-09-448-076-8
7	629	7.7	754	4	US-09-702-572-8
8	622.5	7.6	753	4	US-09-276-400-2
9	622.5	7.6	753	4	US-09-448-076-2
10	622.5	7.6	753	4	US-09-702-572-2
11	607.5	7.5	774	4	US-09-276-400-7
12	607.5	7.5	774	4	US-09-448-076-7
13	607.5	7.5	774	4	US-09-702-572-7
14	462	5.7	100	3	US-09-034-916-15
15	432	5.3	100	3	US-09-034-916-17
16	412	5.1	100	3	US-09-034-916-13
17	347	4.3	100	3	US-09-034-916-11
18	346	4.3	102	3	US-09-034-916-16
19	338	4.2	100	3	US-09-034-916-10
20	337.5	4.1	101	3	US-09-034-916-9
21	326.5	4.0	95	3	US-09-034-916-7
22	311.5	3.8	585	1	US-08-477-674-10
23	311.5	3.8	585	1	US-08-473-791-10
24	311.5	3.8	585	2	US-08-316-714-10
25	311.5	3.8	585	3	US-08-473-673-10
26	311	3.8	100	3	US-09-034-916-14
27	296	3.6	102	3	US-09-034-916-5

28	290.5	3.6	489	2	US-08-794-795-7	Sequence 7, Appl
29	290.5	3.6	489	4	US-09-249-200-7	Sequence 7, Appl
30	290.5	3.6	518	1	US-08-392-367B-2	Sequence 2, Appl
31	290.5	3.6	518	3	US-08-893-467A-2	Sequence 2, Appl
32	290	3.6	451	1	US-08-154-365-2	Sequence 2, Appl
33	288.5	3.5	495	2	US-08-794-795-2	Sequence 2, Appl
34	288.5	3.5	495	4	US-09-249-200-2	Sequence 2, Appl
35	288.5	3.5	520	2	US-08-794-795-6	Sequence 6, Appl
36	288.5	3.5	520	4	US-09-249-200-6	Sequence 6, Appl
37	288	3.5	451	1	US-08-453-117-2	Sequence 2, Appl
38	288	3.5	451	2	US-08-948-222-2	Sequence 2, Appl
39	288	3.5	451	2	US-08-973-145-2	Sequence 2, Appl
40	288	3.5	451	4	US-09-276-400-10	Sequence 10, Appl
41	288	3.5	451	4	US-09-448-076-10	Sequence 10, Appl
42	288	3.5	451	4	US-09-702-572-10	Sequence 10, Appl
43	288	3.5	451	5	PCT-US96-08081-2	Sequence 2, Appl
44	278	3.4	101	4	US-09-518-046-16	Sequence 16, Appl
45	273	3.4	101	4	US-09-518-046-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-09-341-587-3
; Sequence 3, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341.587
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: PCT/DE98/00096
; EARLIER FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1785
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-341-587-3

Query Match	29.9%	Score 2435;	DB 4;	Length 1785;
Best Local Similarity	36.1%	Pred. No. 3.9e-206;		
Matches 533;	Conservative 155;	Mismatches 415;	Indels 374;	Gaps 28;
Qy	43	GTD--LELRVNGDPCSGTVEVKFQGMGTVCDDGNWTTASTVVCQKLGCPFSFAM---	97	
Db	95	GSDSGLALRLVNGDRCQGRVEILYRGSGTVCDDSDNDANVVCRLGCGWANSAPGN	154	
Qy	98	FRFQAVTRHCKIWLDDVSCYGNESALWECOHREWSHNCYHGDEGVNCT-----	148	
Db	155	AWFQGG---SGPIALDDVRCSGHESYLWSCPNGHNLNCHGHDGADGYCISAAQPQSTLR	211	
Qy	149	-----GEANLGLRLVDGNNSCGRVEVKFQERGTCTDCDGNLNTAAV	192	
Db	212	PESWPVRISSPVPTEGSESSALRLVNGDRCRCRVEILYRGSGTVCDDYDWDNDANV	271	
Qy	193	CRLQCGPSSFTISSGVVNSPAVLRPIWLDDILCQGNELALNCRHNGNHDCHSNEDVTL	252	
Db	272	CRLGCGWANSAPGACGQGGPIVLDDVRCSGHESYLWSCPNGHNLNCHGHDGADGY	331	
Qy	253	TCY-----DSSDLRLVGTNRCMGRVELKIQGRWCTVCHHK	290	
Db	332	ICSAQSRPTSPDTPWTPTSHASTAGPSSALRLVNGDRCRCRVEILYRGSGTVCDDSD	391	
Qy	291	WNAADVVCQKLCGCTALHFAGLPHLQSGDVVWLQGVSCSGNESFLWDCRHSGTWNFD	350	
Db	392	WDRSDANVVCRLCGGWATAPGNARFGQSGPIVLDDVRCSGYESYLWSCPNGHNLNCH	451	
Qy	351	CLHQNDYSVICSQDGLLELRDLADGSNNCSGRVEVRIHQWMTICDQNKWNEQALVVCQK	410	

Db 452 CQHSADAGVICSAA-----HSWT-----470
QY 411 GCPFSVFGSRRAPNEARDIWINSICTGNESALWDCTYDGAKAKRTCFRRSDAGVICSD 470
Db 471 -----PSPDLPTITLPAIVGSESS-----491
QY 471 KADLRLVGAHSPCYGRLEVYKQGWGTVCHDRMSTRNAAVVCQKLGCGKPMHVFGMTY 530
Db 492 ---LALRLVGGDRCGRVEVLYQGSWGTVCDDSWDTNDANVVCQPCGGMASAPGNAR 548
QY 531 FKEASGPTLDDVSCIGNESINWDCHEHSGKGNCHVREDVIVTCSGDA-----TW- 581
Db 549 FQGGSGPIVLDVRCSGHESYFWSCPHNGWLSHNCGHSDAGVICSASQSRPTPSPTWP 608
QY 582 -----GURLVGGNRCGRLEVYFOGRWGTVCDDGWNKAAAVVCSOLDPCS 628
Db 609 TSHASTAGSESSALRLVNGDRCGRVEVLYRGSWGTVCDDYDTNDANVVCRLGCGW 668
QY 629 STIAGMLGNASTGYKIWLDDVSCDGDSDLSWCRNSGWNNDCHSHSDEVDGVICS-----683
Db 669 AMSAPGNARFGGSGPIVLDVRCSGHESYFWSCPHNGWLSHNCGHSDAGVICSASQSQ 728
QY 684 -----DASOMELRVGSSRCACGVVNVGGVILCANGWGMNIAEV 726
Db 729 PTPSPDTWPTSHASTAGSESSALRLVNGDRCGRVEVLYRGSWGTVCDDYDTNDANV 788
QY 727 VCRLQECGSAIRVSRPFTERTLHILMSNGCTGEASLWDCIRWENKOTACHLNMESAS 786
Db 789 VCRLQCGWATAPGNARFGGSGPIVLDVRCSGHESYFWSCPHNGWLSHNCGHSDAG 848
QY 787 LICSAHR-OPRLVGDAMPSCGRVEVYKHDATWRSVCDSDFSLHAANVLCRELNCDAISLS 845
Db 849 VICASASQP-----TPSPDTWPT-----867
QY 846 VGDHFGKGNLWAEKFCQEGSETHALCPVQHPEDTCIHSREVVCSRYTDLVNG 905
Db 868 -----SRASTAGSESTLAL-----RLVNG 886
QY 906 KQCGQGOVEINVLHGWSLCTDHPDEARVLCROLSCGTALSTGGKYGIVRSVRVNGH 965
Db 887 GDRGRVLEVLYQGSWGTVCDDYDTNDANVVCRLGCGWASAPGNAQFGGSGPIVLD 946
QY 966 RFHCLGNESLLDNQMTVLGAPPCITGNTVSVICGTSITQ--PLPCLANVSDPYLSAVP 1023
Db 947 DYRCSCGHESYLSWCPHNGWLSHNCGHSDAGVICSASQSRPTDPTLTLNLPALTVGS 1006
QY 1024 EGSALICLEDKRLVLDGSDRCAGRVEIYHDCGFWGTICDDGWLSDAHVVCOKLGGVAF 1083
Db 1007 ESSL-----ALRLVNGDRGRVLEVLYRGSWGTVCDDSWDTNDANVVCRLGCGWAM 1059
QY 1084 NATVSAHFGECSGPIWLDDNLCTGTESHLMQCPRSQGWGQDCHRHEDAGVICSEFTALRL 1143
Db 1060 SAPGNARFGGSGPIVLDVRCSGHESYFWSCPHNGWLSHNCGHSDAGVICS---ATQI 1116
QY 1144 YSET-----ETESACGRLEVY-NGTWGS-----1166
Db 1117 NSTTTDWWHTTTTARPSNCGGL--FYASGTFSPPSYPAYPNNAKCVMEIVNSGY 1174
QY 1167 ---VGRNT-----TTAIGVCI---ROLGCGENGWVSL---AP 1196
Db 1175 RNLGFSNLKLAHNCSDFYVEIFDGSLSNLSLLGKICNDTRQIFTSYNMTHTFRSD 1234
QY 1197 LSKTGSFPM-WYDDIOCPKTHISIWQCLAPWERRISSPAEETWITCEDRIRVRGDTFC 1255
Db 1235 ISFQNTGFLAWN-----SPFSDATL-----RLVNLNLSYGL 1267
QY 1256 SRVETIWHAGSWGTVCDDSDWDLAEAEVVCQQLGCGSALAAALRDASEGQGTGTIWLDDMRC 1315
Db 1268 AGRVEIYHGGTGTVCDDSWTITQAEAEVVCRLGCGRAVSALGNAYFGSGSGPITLDDVBC 1327
QY 1316 KGNESFLWCPHAKPMQSCDGHEDAGVRCSGSLKS 1352
Db 1328 SGTESTLMQCRNGWFSHNCNHRDAGVICSNGHLST 1364

RESULT 2

US-08-470-350B-2
; Sequence 2, Application US/08470350B
; Patent No. 5684126
; GENERAL INFORMATION:
; APPLICANT: Li, Xiao
; APPLICANT: Snyder, Solomon H
; TITLE OF INVENTION: Ebnerin: A Secreted von Ebner's Gland
; TITLE OF INVENTION: Protein Associated with Taste Buds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,350B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Wolfe, Susan A
; REGISTRATION NUMBER: 33,568
; REFERENCE/DOCKET NUMBER: 01107.48790
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1290 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-470-350B-2

Query Match 12.9%; Score 1048.5; DB 1; Length 1290;
Best Local Similarity 29.7%; Pred. No. 1.le-83;
Matches 289; Conservative 114; Mismatches 290; Indels 281; Gaps 34;

QY 286 VCHHKNNAAADVCKQLGCGTALHFAGLPHLQSGSDVVLDGSCSGNESFLWPCRHSQ 345
Db 1 MCDSDNDANVVCRLGCGWALSAPGSAQFGQGLGPIVLDVACRGHEAYLWCSHRG 50
QY 346 TVNFDCLHNDVSVICSDGADLELRLADGNNCSGRVEVRIHQWTTICDQNKNEQALV 405
Db 61 WLSHNCGHEDAGVICSD-----SQTSSPTG-----NW---NPGGTNDVIY 100
QY 406 VCKQLGCPFSVFGSRRAPKPSNEARDIWINSICTGNESALWDCTYDGAKAKRTCFRRSDAG 465
Db 101 DTQET-----TETSQTSSPTDWM-----NHGTTINDVIYD---TOETTEGTDG 142
QY 466 VICSDKADLRLVCAHSPCYGRLEVYKQGWGTVCHDRMSTRNAAVVCQKLGCGKPMH 525
Db 143 -----LAVRLVNGDRGRVLEVLYQGSWGTVCDDSDNDANVVCRLGCGWALS 194
QY 526 FQMTYFKEASGPIWLDDVSCIGNESINWDCHEHSGKGNCHVREDVIVTCS-----576
Db 195 PCSAQFGGSGSIVLDDVACRGHEAYLWCSHRGLSHNCGHEDAGVICSYSQTSSPTP 254
QY 577 -----GDATWG-----LRLVGGNRCGRLEVYFQGRWG 605
Db 255 DSQTSSTPTGWNPNPGTNDVSYGPEQTDTATDSGLAVRLVNGDRCGRVEIYQGSWG 314
QY 606 TVCDDGWNKAAAVVCSOLDPCSIIIGMLGNASTGYKILWLDVSCDGDSDLSWCRNS 665

Db 315 TVCDDSDTKDANVVCRLVCGWALSAPGSAHFQGGSGSIVLDDVACTGHEAYLWSCSHR 374
Qy 666 GWGNDCSHSDVGVICSDAS-----DM-----ELRLVGGSSRCAGKVEVNVQGA 712
Db 375 GWSHNCGHEDAGVICSDAQTSQTPDMPPTTPTTDDWTKYSSVPTQFTI- 433
Qy 713 ILCANGMNTAEVVCRLGECGSAIRVS-----REPHTFERTLHILMSNGCTGGEASLWD 768
Db 434 ---ADMWTPSPPT-----CGLLTLPLPYGQFSSPYPGS-----YPNNARC-----LW- 474
Qy 769 CIRWENKOTACHLANNEASLICSARQPLVCAADPCSGRVEVKHADRWSVCDSDFL- 826
Db 475 -----KIFVSSMNRVTWFTDQLEGG-----CNDYDILVF 505
Qy 827 ----HAANVLCRELNCGDAISLSVDHFGK-----NGLTNAEKFOCE--- 865
Db 506 DGPENNSLIAR-----VCDGFNGSFTSTQNFMSVVFITDGSVTRRGQADYYS 554
Qy 866 ----GSETHALCPIV-----QHPEDTG-----IHSREVGVCSRYTDVRLVNGKSQCQGV 913
Db 555 TPISTSTPTTPTTPIVTDWMTTPSPVTCGGLLTLPLPYGQFSSPY----- 598
Qy 914 EINVLGHWSLCDTHWDPEDARVLCR-----QLSCGT-----ALSTTGK 953
Db 599 -----YPGSY-----PNNARCWKIFVPSMNRVTWFTDQLEGGCNDYILGFDGPE 646
Qy 954 YIGERSVRVGMHRFCHLGNESLLDNCOMTVLGAAPCIHNTVSV---ICTGSLTOPLF--- 1008
Db 647 YNSSLIARV-----CDGNSGFTSTQ-----NFMVVFITDGSVTRRGFOAD 688
Qy 1009 ----PCLANSDPYLSAPEG---SALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICD.1062
Db 689 YYSTPIRTSTPTTPTTPIITGNDSSLV-----LRLVNGTNRCEGRVEILYRGSWVPACD 742
Qy 1063 DGWLDSDAHVVCQKLGCGVARNATVSAHFEGGSGPIWDLNCTGTESHLMQCPGSRGWGQ 1122
Db 743 DSWINDANVVCRLGCGSALSAPGNAWFGGSLVILVDVYSCGYESHLCNRHPGLV 802
Qy 1123 HCRKHEDAGVICS 1136
Db 803 HNCRVEDAGVICS 816

RESULT 3

US-09-341-587-1
; Sequence 1, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341,587
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: PCT/DE98/00096
; EARLIER FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-341-587-1

Query Match 8.88; Score 718.5; DB 4; Length 666;
Best Local Similarity 35.88; Pred. No. 6.7e-55;
Matches 162; Conservative 63; Mismatches 118; Indels 109; Gaps 13;
Qy 971 GNESLLDNCOMTVLGAAPCIHNTVSVICTGSLTQ--PLFPCLANVSDPYLSAVPEGSAL 1028
Db 1 GHESVLSNCPHNGWLSHNCGHEDAGVICSAAQSQSTPRPTLTTLNLPALTVGSESL- 59
Qy 1029 ICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVVCQKLGCGVAFNATVS 1088
|||||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:|

Db 60 -----ALRLVNGDRCRGRVEVLYRGSWGTVCDDSDWTDNDANVVCRLGCGWAMSAPGN 113
Qy 1089 AHFEGGSGPIWDLNCTGTESHLMQCPGSRGWGDCHRKEDAGVICSEFTALRYSET- 1147
Db 114 ARFGGSGPIWDDVRCSGNESYILWSCPHKGLWLNCHGHHEDAGVICS---ATQINSTTT 170
Qy 1148 -----ETESCAGRLVEFY-NGTWGS-----VG 1168
Db 171 DMWHPTTTTTPARPSSNCGFL--FYASGTFSSPSYPAYYPNNAKVMEIVNSGVYRNLG 228
Qy 1169 RRNI-----TTAIGIVC-----RQLCGGNGVYSL---APLSKTG 1201
Db 229 FSNLKEAHNCSFDYVEIFDGLNSLLGKICNDTRQIFTSSYNRMTIHFSDISFQN 288
Qy 1202 SGFM-WVDDIOCPKTHISIWQLSAPWERRISSPAEETWITCEDRIRVRGGDTSCGRVE 1260
Db 289 TGFLAWN-----SFFSDATL-----RLVNLNSSYGLCAGRVE 321
Qy 1261 IWHAGSMGTVCDDSDWDLAEAEVVCQQLCGSALALRDASFQGGTGTIWLDDMRCKGNES 1320
Db 322 IYHGTGWTVCDDSWTIQEAEBVVCRLCGRAVSALGNAYFGSGGPIITLDDVECSGTES 381
Qy 1321 FLWDCHAKPWQSCDGHKEDAGVRCSGSLAS 1352
Db 382 TLWQCRNRGWFSHNCNHRDAGVICSGNHLST 413

RESULT 4

US-09-034-916-2
; Sequence 2, Application US/09034916
; Patent No. 6046314
; GENERAL INFORMATION:
; APPLICANT: GEBE, JOHN A.
; APPLICANT: SIADAK, ANTHONY W.
; APPLICANT: ARUFFO, ALEJANDRO A.
; TITLE OF INVENTION: SPALPHA: A NOVEL SCAVENGER RECEPTOR
; TITLE OF INVENTION: CYSTEINE-RICH DOMAIN-CONTAINING POLYPEPTIDE AND MONOCLONAL
; TITLE OF INVENTION: ANTIBODIES THERETO
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBINS & ASSOCIATES
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/034,916
; FILING DATE: 04-MAR-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,956
; FILING DATE: 06-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 5998-0020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 325-7812
; TELEFAX: (650) 325-7823
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 346 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-034-916-2

```

Query Match      7.7%; Score 629; DB 4; Length 754;
Best Local Similarity 31.0%; Pred. No. 6.7e-47;
Matches 155; Conservative 72; Mismatches 183; Indels 90; Gaps 16;

QY      56 PCSTVEVKFGQWGTTCDDGWNNTASTVYVCKQLGCFPSFA-----MRFQCAVTRHGKI 110
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db      54 PYGRVEIQRAGWGTTICDDDTFTLQAHLVLCRELG--FTEATGHTSAKYCGPT---GRI 108
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |

QY      111 WLDDVSCYGNESALWECQHREWSGNSHCYVGBDVGVC-----YGEANL----- 153
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db      109 WLDNLSCRGTEGSGYECASRGSGNSDCTHDDAGVCKDQLPGFSDSNVLEVBHQLOVE 168
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |

QY      154 -----GLRLVDGNNCSGRVEYKFQERWGTICDDGNLNTAAVYVQLGCPSS-- 201
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db      169 EVRLRPVAVGRRPLP---VTEGLVEVRLPEGWSQVCDKGWSAHNSHVVCMLGFPPEKR 225
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |

QY      202 -----FISSGVVNSPAVLRPIMLDDILCOGNELANCRHRRGWNHD---CSHNEDVT 251
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db      226 VNMAFYRLAOKQHS-----FGLHSVACYGTGTEAHLSLCSLEFYRANDTTRCSGNGPAV 279
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |

QY      252 LTCY-----DSSDLELRVLGGTNRCMGKRVELKIQGRWGTVCVCHKKWN 293
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db      280 VSCVLGPLYATFTQCKKQKHKPGQEARVKRUKGAHQGEGRVEVLKAGTWGTVCDKRKDWL 339
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |

QY      294 AAADVCKQLGCGTALHFAGLPHLQSGSDVYWLVDGVS CSGNESFTLWDCRHSGT VNFDC 353

```

[illegible]

TYPE: PRT
ORGANISM: Homo sapiens
US-09-448-076-2

Query Match
Best Local Similarity 7.6%; Score 622.5; DB 4; Length 753;
Matches 167; Conservative 71; Mismatches 202; Indels 95; Gaps 19;

QY 30 LLLNSCLFSSFNPTG-----DLELRVN-GDGPCSGTVEVKFGQGWGTVCD 74
DB 14 LLL--CLLSSCLGSPSTGPEKKAGSQGLRFRLAGPRKPYEGRVEIQRAGWGTCID 71
QY 75 DGNWTTASTVVKQKGCFFSFA-----MFRFGQAVTRHGKIWLDDVSCYGNESALWECOH 129
DB 72 DDFTLQAAHILCRELG--FTEATGTHSAKYGPGT--GRIWLDNLSCGTEQSVTECAS 126
QY 130 REMGSHNCHGEDYGVNC-----YGEANL-----GLRLVDGNN 162
DB 127 RGWNSDCTHDEDAVICKDQRLPGFSDSNVIEVEHHIQVEEVRIRPAVGWRPLP--- 183
QY 163 SCGRVEVKFQERWGTICDDGNLNTAAVVCRLGCPSPSFSSGVVNSPAVL---PIWL 219
DB 184 VTEGLVEVRLPDGWSQVCDKWSAHNSHVCGMLGFPSEKRVNAAFYRLLAQROHSGFL 243
QY 220 DDILCOGNELALWNCRRHGWGNDH--CSHNEDVTLC-----YDSS----- 258
DB 244 HGVACVGTGEAHLSCSLEFYRANDTARCPGGPVPVYAASSGQKKQOQSKPQG 303
QY 259 DLELRVCGTNRGMRVELKIOGRWGTVCHHKWNNAAADVCKOLGCGTALHFAGLPHLO 318
DB 304 EARVRLKGAHPGEGRVEVLKASTGTGTCDDRWDLHAASVVCRELFGSAREALSGARMG 363
QY 319 SGSDVVWLDGVSCGNEFSLWDCRHSHTVNFDCLEHNDVSVICS---DGADLELRADGS 375
DB 364 QMGAIHLSEVRCSEGLSLMKCPKHNITAEDCSHQDAGVRCNLPYTGAEIRLSGR 423
QY 376 NNCGRVEVRIHE-----QWTTICDNWKNQALVVKOLGCPFPFVSGSRR--KPSNEAR 429
DB 424 SQHEGRVEVQIGPGPLRWGLTCGDDWGTLEAMVACROGLGLGYANHGLQETWYDWSGNT 483
QY 430 DIWINSICTGNESALWDCYDGKAKRTRCFR--RSDAGVICSDDKADLDLRLVGA 481
DB 484 EYVMSGVRCTGTELSLQCAHHG--THITCKRTGTRTAGVICSETAS-DLLLHSA 536

RESULT 10
US-09-702-572-2
; Sequence 2, Application US/09702572
; Patent No. 6391602
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
; FILE REFERENCE: MNI-073
; CURRENT APPLICATION NUMBER: US/09/702,572
; PRIOR FILING DATE: 2000-10-31
; PRIOR FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 753
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-702-572-2

Query Match
Best Local Similarity 7.6%; Score 622.5; DB 4; Length 753;
Matches 167; Conservative 71; Mismatches 202; Indels 95; Gaps 19;

QY 30 LLLNSCLFSSFNPTG-----DLELRVN-GDGPCSGTVEVKFGQGWGTVCD 74
DB 14 LLL--CLLSSCLGSPSTGPEKKAGSQGLRFRLAGPRKPYEGRVEIQRAGWGTCID 71

QY 75 DGNWTTASTVVKQKGCFFSFA-----MFRFGQAVTRHGKIWLDDVSCYGNESALWECOH 129
DB 72 DDFTLQAAHILCRELG--FTEATGTHSAKYGPGT--GRIWLDNLSCGTEQSVTECAS 126
QY 130 REMGSHNCHGEDYGVNC-----YGEANL-----GLRLVDGNN 162
DB 127 RGWNSDCTHDEDAVICKDQRLPGFSDSNVIEVEHHIQVEEVRIRPAVGWRPLP--- 183
QY 163 SCGRVEVKFQERWGTICDDGNLNTAAVVCRLGCPSPSFSSGVVNSPAVL---PIWL 219
DB 184 VTEGLVEVRLPDGWSQVCDKWSAHNSHVCGMLGFPSEKRVNAAFYRLLAQROHSGFL 243
QY 220 DDILCOGNELALWNCRRHGWGNDH--CSHNEDVTLC-----YDSS----- 258
DB 244 HGVACVGTGEAHLSCSLEFYRANDTARCPGGPVPVYAASSGQKKQOQSKPQG 303
QY 259 DLELRVCGTNRGMRVELKIOGRWGTVCHHKWNNAAADVCKOLGCGTALHFAGLPHLO 318
DB 304 EARVRLKGAHPGEGRVEVLKASTGTGTCDDRWDLHAASVVCRELFGSAREALSGARMG 363
QY 319 SGSDVVWLDGVSCGNEFSLWDCRHSHTVNFDCLEHNDVSVICS---DGADLELRADGS 375
DB 364 QMGAIHLSEVRCSEGLSLMKCPKHNITAEDCSHQDAGVRCNLPYTGAEIRLSGR 423
QY 376 NNCGRVEVRIHE-----QWTTICDNWKNQALVVKOLGCPFPFVSGSRR--KPSNEAR 429
DB 424 SQHEGRVEVQIGPGPLRWGLTCGDDWGTLEAMVACROGLGLGYANHGLQETWYDWSGNT 483
QY 430 DIWINSICTGNESALWDCYDGKAKRTRCFR--RSDAGVICSDDKADLDLRLVGA 481
DB 484 EYVMSGVRCTGTELSLQCAHHG--THITCKRTGTRTAGVICSETAS-DLLLHSA 536

RESULT 11
US-09-276-400-7
; Sequence 7, Application US/09276400
; Patent No. 6140056
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
; FILE REFERENCE: MNI-073
; CURRENT APPLICATION NUMBER: US/09/276,400
; CURRENT FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 7
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-276-400-7

Query Match
Best Local Similarity 7.5%; Score 607.5; DB 4; Length 774;
Matches 196; Conservative 86; Mismatches 258; Indels 163; Gaps 28;

QY 260 LELRLVGGTNR-CMGRVELKIQGRWGTVCHHKWNNAAADVCKOLGCGTALHFAGLPHLO 318
DB 56 IOLRLAGOKRKHSEGRVEVYDQGWGTVCDDDFSIHAAHVVCRELGYVEAKSWTASSYG 115
QY 319 SGSDVVWLDGVSCGNEFSLWDCRHSHTVNFDCLEHNDVSVICS---DGADLELRADGS 374
DB 116 KEGEPIWLDNLHCTGNEATLAACTNSGNGWGTCDKTEEDGVVVCSDKRIPIGFKEDNSLNG 175
QY 375 SNNS-----GRVEVRTHEOWTICDNWKNQALVVKQKLG 411
DB 176 IENLNIQVEDIRAILSTYKRTPYMEGYVEVEKGTWKQICDKHWAKNSRVVCGMFG 235
QY 412 CP-----FSVFGSRRRAKPSNEARDIWNISICTGNESALWDC-----TYDGAKR 456
DB 236 FPGERTYNTKVYKMFASRR-----KQRYWPFSDCTGTGTEAHISSCKLGPQVSLDPMKNV 289

Db 236 PFERTYNTKVKMFASRR-----KQYWPFSMDCTEAISSCKLGPQVSLDPMKNV 289
QY 457 TCFRSDAGVIC-----SDKADLDL-RLVGAHSPCYGRLEVKYQGEWGT 499
Db 290 TCENGLPAVSCVPCQVSPDGPSPFRKAYKPEQPLRLRGAYIGEGRVEVLKNGEWT 349
QY 500 VCHDRWSTRNAAVCKQLCGCKPMHVFMTVFKEASGPIWLLDDVSCIGNESNIWDCESG 559
Db 350 VCDKDWLVSASVCRLELFGSAAKAVTCSRLGQIGIPHLNEIOCTGNEKSLIDCKENA 409
QY 560 WGHKNCVHREDVIVTCSDATW---GLRVGSGNRCGRLEVIYFQGR---WGTVCDDGW 612
Db 410 -ESQCNHEEDAGVRCNTPAMGLQKRLNGRNPYEGRVEVLVERNGSLVMGMVCGQNW 468
QY 613 NSKAAVVCSQLDCPSSIIIGMGLNASTGYGKIW-----LDDVSCDGDSDILW 660
Db 469 GIVEAMVVCRL-----GLGFASNAFOETWTHGVDVSNKVMVSGVCKSGTELSLA 519
QY 661 SCRNWGNNDGSHSEV-----GVICSD-ASDMEL--RLVGGSSRCAGKVEVN 706
Db 520 HCRHDG-----EDVACPOGGVQYGAGVACSETAPDLVLAEMVQOTIYLEDPRMFM 570
QY 707 VQGAIVGILCANGWMNIAEVCWRCQLECGSAIRVSRPHFTERTLHILMS-----NSGCT--- 760
Db 571 LQ-----CA-----MEENCLASASAAQTDPTTGYRRLRFRSQIHNNQSDFR 612
QY 761 ---GGEASLW-DCIR-WEWKQOTACH---LNMEASLICSNAHROPRLVGADPCSGRVEVKH 812
Db 613 PKNGRHWIWHDCRRHYSMEVFTHYDLLNLNGTKVAEGHKASFCLEDECEGDIQXNY 671
QY 813 ADTWSVC-----DSOFSLHAANVLCRELNGC--DAISLSVGDH 849
Db 672 E-----CANFGDQGITNCMDMYRHDIDCQWVDITDVPBGDY 708

RESULT 14
US-09-034-916-15
; Sequence 15, Application US/09034916
; Patent No. 6046314
; GENERAL INFORMATION:
; APPLICANT: GEBE, JOHN A.
; APPLICANT: SIADAK, ANTHONY W.
; APPLICANT: ARUFFO, ALEJANDRO A.
; TITLE OF INVENTION: SPALPHA: A NOVEL SCAVENGER RECEPTOR
; TITLE OF INVENTION: CYSTEINE-RICH DOMAIN-CONTAINING POLYPEPTIDE AND MONOCLONAL
; TITLE OF INVENTION: ANTIBODIES THERETO
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBINS & ASSOCIATES
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/034,916
; FILING DATE: 04-MAR-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,956
; FILING DATE: 06-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 5998-0020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 325-7812
; TELEFAX: (650) 325-7823
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-034-916-17

Query Match 5.3%; Score 432; DB 3; Length 100;
Best Local Similarity 71.0%; Pred. No. 8.7e-31;
Matches 71; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

TELEFAX: (650) 325-7823
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-034-916-15

Query Match 5.7%; Score 462; DB 3; Length 100;
Best Local Similarity 77.0%; Pred. No. 1.9e-33;
Matches 77; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1037 RLVDGDSRCAGRVEIYHDGFWGTTCDDGWLSDAHVVCQKLCGCVAFNATYSAHFGECSG 1096
Db 1 RLVNGGGRCAGRVEIYHSGSWGTCDDSDWLSDAHVVCRLGCGEAINATGSAHFGECSG 60
QY 1097 PIWLDLCTGTESHLWOCPSRGWGQCHDKHEDAGVICS 1136
Db 61 PIWLDKMGKESRIMQCHSHGWGQQNCRHKEDAGVICS 100

RESULT 15
US-09-034-916-17
; Sequence 17, Application US/09034916
; Patent No. 6046314
; GENERAL INFORMATION:
; APPLICANT: GEBE, JOHN A.
; APPLICANT: SIADAK, ANTHONY W.
; APPLICANT: ARUFFO, ALEJANDRO A.
; TITLE OF INVENTION: SPALPHA: A NOVEL SCAVENGER RECEPTOR
; TITLE OF INVENTION: CYSTEINE-RICH DOMAIN-CONTAINING POLYPEPTIDE AND MONOCLONAL
; TITLE OF INVENTION: ANTIBODIES THERETO
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBINS & ASSOCIATES
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/034,916
; FILING DATE: 04-MAR-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,956
; FILING DATE: 06-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 5998-0020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 325-7812
; TELEFAX: (650) 325-7823
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-034-916-17

Query Match 5.3%; Score 432; DB 3; Length 100;
Best Local Similarity 71.0%; Pred. No. 8.7e-31;
Matches 71; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

QY 181 DDGWLNTAAVVCRLGCPSPSEISSGVNNSPAVLRIPIWDDILCOGNELALWNCRRHWG 240
Db 181 DDGWLNTAAVVCRLGCPSPSEISSGVNNSPAVLRIPIWDDILCOGNELALWNCRRHWG 240
QY 241 NHDCHSHNEDVLTTCYDSDLELRLVGGTNRGCMGRVELKIQGRWGTVCCHKKNNAAADVVC 300
Db 241 NHDCHSHNEDVLTTCYDSDLELRLVGGTNRGCMGRVELKIQGRWGTVCCHKKNNAAADVVC 300
QY 301 KOLGCGTALHFAGLPHLGSGSDVWLDGVSGNSFSLWDCRHSFTVNFDCIHLQNDVSVI 360
Db 301 KOLGCGTALHFAGLPHLGSGSDVWLDGVSGNSFSLWDCRHSFTVNFDCIHLQNDVSVI 360
QY 361 CSDGADLELRLADGNNCGRVEYRIHQWMTICDNKWNKNEALVVKOLGCPFSVFGSR 420
Db 361 CSDGADLELRLADGNNCGRVEYRIHQWMTICDNKWNKNEALVVKOLGCPFSVFGSR 420
QY 421 RAKPSNEARDIWNISCTGNSALWDCYDCKAKRTCFRSDAGVICSADKADLRLVG 480
Db 421 RAKPSNEARDIWNISCTGNSALWDCYDCKAKRTCFRSDAGVICSADKADLRLVG 480
QY 481 AHSPCYGRLEVYQGEWGTVCCHDRSTRNAAVVCOLGCGKPMHVFGMTYFKEASGPIWL 540
Db 481 AHSPCYGRLEVYQGEWGTVCCHDRSTRNAAVVCOLGCGKPMHVFGMTYFKEASGPIWL 540
QY 541 DDVSCIGNESNIWDCEHSGWGHKNCVHREDVITVCSGDATWGLRLVGGSNRCSGRLEYF 600
Db 541 DDVSCIGNESNIWDCEHSGWGHKNCVHREDVITVCSGDATWGLRLVGGSNRCSGRLEYF 600
QY 601 QGRWGTVCDDGWNKAAAVVCVQSDCPSSIIGMGLGNASTGYKTIWDDVSCDGEDSLW 660
Db 601 QGRWGTVCDDGWNKAAAVVCVQSDCPSSIIGMGLGNASTGYKTIWDDVSCDGEDSLW 660
QY 661 SCRNSGWNDCSHSEDEVGVICSDASDMELRLVGGSSRCAGVEVNOGAVGILCANGWG 720
Db 661 SCRNSGWNDCSHSEDEVGVICSDASDMELRLVGGSSRCAGVEVNOGAVGILCANGWG 720
QY 721 MNAEVCVQLCGGSAIRVSRPHFTERTLHLILMNSGCTGGEASLWDCIRWEWQTACH 780
Db 721 MNAEVCVQLCGGSAIRVSRPHFTERTLHLILMNSGCTGGEASLWDCIRWEWQTACH 780
QY 781 LNEASLICSASRQRLVGCADMPGSRVEVKHADTWRSVCDSDFSIHAANVLCRELNGD 840
Db 781 LNEASLICSASRQRLVGCADMPGSRVEVKHADTWRSVCDSDFSIHAANVLCRELNGD 840
QY 841 AISLSYGDHFGKNGLTWAEKFCQEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTIV 900
Db 841 AISLSYGDHFGKNGLTWAEKFCQEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTIV 900
QY 901 RLVNGSKQDQGVIEINVLGHWGLSCDTHWDPEDARVLCRLSCGTALSTTGKYGIRSV 960
Db 901 RLVNGSKQDQGVIEINVLGHWGLSCDTHWDPEDARVLCRLSCGTALSTTGKYGIRSV 960
QY 961 RVWGHFCHLGNESLDNCOMTVLAPPCIHGNTVSVICTGSLTOPLPCLANVSDPYLS 1020
Db 961 RVWGHFCHLGNESLDNCOMTVLAPPCIHGNTVSVICTGSLTOPLPCLANVSDPYLS 1020
QY 1021 AVPEGSALICLEDKRLRLVGDGSRACGRVEIYHDFGFWGTICDDGWDLSDAHVVVCKLGGC 1080
Db 1021 AVPEGSALICLEDKRLRLVGDGSRACGRVEIYHDFGFWGTICDDGWDLSDAHVVVCKLGGC 1080
QY 1081 VAFNATVSAHFEGSGPIWLDLNLCTGTESHLWQCPSRGWGQHDCHRHEDAGVICSFTA 1140
Db 1081 VAFNATVSAHFEGSGPIWLDLNLCTGTESHLWQCPSRGWGQHDCHRHEDAGVICSFTA 1140
QY 1141 LRLYSETETESACRLEVEFYNGTWGSGVRNNTTATAGTVCRQLGCGENGVS LAPLSKT 1200
Db 1141 LRLYSETETESACRLEVEFYNGTWGSGVRNNTTATAGTVCRQLGCGENGVS LAPLSKT 1200
QY 1201 GSGFMVDDIQCPKTHISTWQCLSAFWERRISSPAETWITCEDRIRVRGGDTECSGRVE 1260
Db 1201 GSGFMVDDIQCPKTHISTWQCLSAFWERRISSPAETWITCEDRIRVRGGDTECSGRVE 1260
QY 1261 IWHAGSWGTVCDSDWDLAEAEVVCQQLGCGSALALRDSAFQCGGTIWLDDMRCKGNES 1320

Db 1261 IWHAGSWGTVCDSDWDLAEAEVVCQQLGCGSALALRDSAFQCGGTIWLDDMRCKGNES 1320
QY 1321 FLWCHAKPWQSCDGHKEDAGVRCSGQSLKSNASSGHLALILSSIFGLLLVLFILFL 1380
Db 1321 FLWCHAKPWQSCDGHKEDAGVRCSGQSLKSNASSGHLALILSSIFGLLLVLFILFL 1380
QY 1381 TWCVRQKQKHLPLRVSTRRRGSLLENLFHEMETCLKREDPHGTRTSDTPNHGCEASDT 1440
Db 1381 TWCVRQKQKHLPLRVSTRRRGSLLENLFHEMETCLKREDPHGTRTSDTPNHGCEASDT 1440
QY 1441 SLLGVLPASEATK 1453
Db 1441 SLLGVLPASEATK 1453
RESULT 2
US-09-759-130B-381
; Sequence 381, Application US/09759130B
; Publication No. US2003002279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: MP100-5350MIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 381
; LENGTH: 1453
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-130B-381
Query Match 100.0%; Score 8138; DB 9; Length 1453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WMLPNSWHIDGRCCCHQNLFSVAVVTCILLNSCFLLISSFNGTDLRLVNGDPCSCGT 60


```
Db 1 MMLPQNSWHIDFRGCCCHQNLFSAVVTCILLNSCFLISSFNCTDLELRVLVNGDPCSGT 60
QY 61 VEYKFGQMGTCVDDGWNNTASTVVCVKQLGCPSEFAFRGQAVTRHGKWLDDVSCYGN 120
Db 61 VEYKFGQMGTCVDDGWNNTASTVVCVKQLGCPSEFAFRGQAVTRHGKWLDDVSCYGN 120
QY 121 ESALWECQHQREMGSHNCHYHGEDYGVNCGYGANLGLRLVDGNNCSGRVEVKFOERMGTC 180
Db 121 ESALWECQHQREMGSHNCHYHGEDYGVNCGYGANLGLRLVDGNNCSGRVEVKFOERMGTC 180
QY 181 DDGWNLNTAAVVCRLQGCPSFSSTSSGVVNSPAVLRLPWLDDIILCOGNELALNCRHGWG 240
Db 181 DDGWNLNTAAVVCRLQGCPSFSSTSSGVVNSPAVLRLPWLDDIILCOGNELALNCRHGWG 240
QY 241 NHDCSHNEDVTLTCYDSSDLELRVLGVTNRCMGRVELKIQGRMGTCVCHHKWNNAADVVC 300
Db 241 NHDCSHNEDVTLTCYDSSDLELRVLGVTNRCMGRVELKIQGRMGTCVCHHKWNNAADVVC 300
QY 301 KQLGCGTALHFAGLPHLQSGSDVWLDGVSCSGNESFLWDCRHSRGTVNFDCLRHQNDVSVI 360
Db 301 KQLGCGTALHFAGLPHLQSGSDVWLDGVSCSGNESFLWDCRHSRGTVNFDCLRHQNDVSVI 360
QY 361 CSDGADLELRADGNNCSGRVEVRHEQWNTICDQWKNNEQALVYCKQLGCPSEFVGSR 420
Db 361 CSDGADLELRADGNNCSGRVEVRHEQWNTICDQWKNNEQALVYCKQLGCPSEFVGSR 420
QY 421 RAKPSNEARDIWNISCTGNEALMDCTYDGRKRTCFRRSDAGVTCSDKADLDLRLVG 480
Db 421 RAKPSNEARDIWNISCTGNEALMDCTYDGRKRTCFRRSDAGVTCSDKADLDLRLVG 480
QY 481 AHSPCYGRLEVKYQGEWGTVCHDRWSTRNAVVCVKQLGCGKPMHVFGMTYFKEASGPWL 540
Db 481 AHSPCYGRLEVKYQGEWGTVCHDRWSTRNAVVCVKQLGCGKPMHVFGMTYFKEASGPWL 540
QY 541 DVVSCIGNESINWDCRHSKGKNCVHREDVITVCSGDATWGLRLVGGNRCSGRLEVYF 600
Db 541 DVVSCIGNESINWDCRHSKGKNCVHREDVITVCSGDATWGLRLVGGNRCSGRLEVYF 600
QY 601 QGRWGTCDGWNKAAVVCSDQPCSSITIGMGLNASTGYGKIWLDDVSCDGDDESILW 660
Db 601 QGRWGTCDGWNKAAVVCSDQPCSSITIGMGLNASTGYGKIWLDDVSCDGDDESILW 660
QY 661 SCRNSGWNNDSCSHSEDDVGVICSDASDMLRLVGGSSRCAGKVEVNVQVAGILCANGWG 720
Db 661 SCRNSGWNNDSCSHSEDDVGVICSDASDMLRLVGGSSRCAGKVEVNVQVAGILCANGWG 720
QY 721 MNTAEVVCRLQCGSALRVSRPHFTERTLHILMSNGSGCTGGEASLWDCIRWEKQTACH 780
Db 721 MNTAEVVCRLQCGSALRVSRPHFTERTLHILMSNGSGCTGGEASLWDCIRWEKQTACH 780
QY 781 LNMEASLICSARHQPRLVGMPCSGRVEVKHADTWRSVCDPSDFSLHAANVLCRELNCGD 840
Db 781 LNMEASLICSARHQPRLVGMPCSGRVEVKHADTWRSVCDPSDFSLHAANVLCRELNCGD 840
QY 841 AISLSVGDHFGKNGLTWAEKFCQEGSETHALCPIVQHPEDTCIHSREYGVVCSRYTDV 900
Db 841 AISLSVGDHFGKNGLTWAEKFCQEGSETHALCPIVQHPEDTCIHSREYGVVCSRYTDV 900
QY 901 RLNVGKSCQDQGVIEINVLGHWGSLCDTHWDPEDARVLCRLSCGTALSTTGGKYIGERSV 960
Db 901 RLNVGKSCQDQGVIEINVLGHWGSLCDTHWDPEDARVLCRLSCGTALSTTGGKYIGERSV 960
QY 961 RVGHRPHCLGNESLNDNCQMTVLGAPPCIRHNTVSVICTGSLTOPFLPCLANVSDPYLS 1020
Db 961 RVGHRPHCLGNESLNDNCQMTVLGAPPCIRHNTVSVICTGSLTOPFLPCLANVSDPYLS 1020
QY 1021 AVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTCICDGDWDLSDAHVVCQKLGCG 1080
Db 1021 AVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTCICDGDWDLSDAHVVCQKLGCG 1080
QY 1081 VAFNATVSAHFEGSGPWLDDNLCTGTESHLWQCPSPRGWGQHDCHRIKEDAGVICSFTFA 1140
Db 1081 VAFNATVSAHFEGSGPWLDDNLCTGTESHLWQCPSPRGWGQHDCHRIKEDAGVICSFTFA 1140
```

```
QY 1141 LRLYSETETESACGRLEVFYNGTWGSGVRNRNITTATAGIVCRLGCGENGWVSLAPLSKT 1200
Db 1141 LRLYSETETESACGRLEVFYNGTWGSGVRNRNITTATAGIVCRLGCGENGWVSLAPLSKT 1200
QY 1201 GSGFMWDDIQCCKTHISIWQCLISAPWERRISSPABETWITCEDRIRVRGDTGTECSRVE 1260
Db 1201 GSGFMWDDIQCCKTHISIWQCLISAPWERRISSPABETWITCEDRIRVRGDTGTECSRVE 1260
QY 1261 IWHAGSWGTVCDSDWDLAEAEVVCQOLGCGSALAAALRDASFQGGTGTIWLDDMRCKGNES 1320
Db 1261 IWHAGSWGTVCDSDWDLAEAEVVCQOLGCGSALAAALRDASFQGGTGTIWLDDMRCKGNES 1320
QY 1321 FLWDCHAKPMWQSDCCGHKEDAGVRCQSLKSNASSGHLALILSSIFGLLLVLFILFL 1380
Db 1321 FLWDCHAKPMWQSDCCGHKEDAGVRCQSLKSNASSGHLALILSSIFGLLLVLFILFL 1380
QY 1381 TWCVRQKQKHLPLRVSTRRRRGSLEENLFHEMETCLKREDPHGTTRTSDTPNHGCEADSDT 1440
Db 1381 TWCVRQKQKHLPLRVSTRRRRGSLEENLFHEMETCLKREDPHGTTRTSDTPNHGCEADSDT 1440
QY 1441 SLGVLPASEATK 1453
Db 1441 SLGVLPASEATK 1453
```

RESULT 3

```
US-10-042-431-13
; Sequence 13, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: PRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-602
; CURRENT APPLICATION NUMBER: US/10/042,431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1413
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-431-13
```

Query Match 97.2%; Score 7913; DB 9; Length 1413;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 41 FNGTDLRLVLVNGDPCSGTVVEVKFQGWGTVCDDGNNTTASTVVCVKQLGCPSEFAMFRF 100
Db 1 FNGTDLRLVLVNGDPCSGTVVEVKFQGWGTVCDDGNNTTASTVVCVKQLGCPSEFAMFRF 60
QY 101 GQAVTRHGKIWLDDVSCYGNESALWECQHQREMGSHNCHYHGEDVGVNCGYGANLGLRLVDG 160
Db 61 GQAVTRHGKIWLDDVSCYGNESALWECQHQREMGSHNCHYHGEDVGVNCGYGANLGLRLVDG 120
QY 161 NNSCSGRVEVKFOERMGTCICDGNLNTAAVVCRLQGCPSFSSTSSGVVNSPAVLRLPWL 220
Db 121 NNSCSGRVEVKFOERMGTCICDGNLNTAAVVCRLQGCPSFSSTSSGVVNSPAVLRLPWL 180
QY 221 DILCOGNELALNCRHGWGNHDCSHNEDVTLTCYDSSDLELRVLGVTNRCMGRVELKIQ 280
Db 181 DILCOGNELALNCRHGWGNHDCSHNEDVTLTCYDSSDLELRVLGVTNRCMGRVELKIQ 240
QY 281 GRMGTCHHKWNNAADVVCVKQLGCGTALHFAGLPHLQSGSDVWLDGVSCSGNESFLWD 340
```

Db 241 GRWGTVCCHKWNNAADVYVQKCGTALHFAGLPHLGSGDVVWLDGVCSGNESFLWD 300
QY 341 CHSGTAVNFDCLHONDVSVICSDGADLELRLADGNNCSGRVEVRIHEOWNTICDONWKN 400
Db 301 CHSGTAVNFDCLHONDVSVICSDGADLELRLADGNNCSGRVEVRIHEOWNTICDONWKN 360
QY 401 EQALVVVQKLGCPFSVFGSRRAPKSNARDIWINISICTGNFESALWDCYDGGKAKRTCFR 460
Db 361 EQALVVVQKLGCPFSVFGSRRAPKSNARDIWINISICTGNFESALWDCYDGGKAKRTCFR 420
QY 461 RSDAGVICSADKADLRLVGAHSPCYGRLVYKQEGWTVCHDRWSTRNAAVVCKQLCGG 520
Db 421 RSDAGVICSADKADLRLVGAHSPCYGRLVYKQEGWTVCHDRWSTRNAAVVCKQLCGG 480
QY 521 KPMHVFGMTYKPEASGPTWDDVSCIGNESNIWDCHEHSGWKHNCVHREDVIVTCSGAT 580
Db 481 KPMHVFGMTYKPEASGPTWDDVSCIGNESNIWDCHEHSGWKHNCVHREDVIVTCSGAT 540
QY 581 WGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNKSAAAVVCQSDPCSSIICMGLGNAST 640
Db 541 WGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNKSAAAVVCQSDPCSSIICMGLGNAST 600
QY 641 GYGTWLDVSCDGEDSLWSCRNSGWNNDCHSHEDVGVICSDASDMELRLVGGSSRCA 700
Db 601 GYGTWLDVSCDGEDSLWSCRNSGWNNDCHSHEDVGVICSDASDMELRLVGGSSRCA 660
QY 701 GKVEVNVQAGVILCANGWGMIAEVVCRQLPCGSAIRVSRPHETERTLHILMNSGCT 760
Db 661 GKVEVNVQAGVILCANGWGMIAEVVCRQLPCGSAIRVSRPHETERTLHILMNSGCT 720
QY 761 GGEASLWDCIRWEKQTAACHLWNEASLICSARHQPRLVGMADPCSGRVEVHADTWRSVC 820
Db 721 GGEASLWDCIRWEKQTAACHLWNEASLICSARHQPRLVGMADPCSGRVEVHADTWRSVC 780
QY 821 DSDFSLHAANVLCRELNCDAISLVGDHFGKNGLTWAEKFCQEGSETHALCPVIOHP 880
Db 781 DSDFSLHAANVLCRELNCDAISLVGDHFGKNGLTWAEKFCQEGSETHALCPVIOHP 840
QY 881 EDTCHSRREVGVVCSRYTDLVRLVNGKSCDGOVEINVLGHWSLCTDTHWDPEDARVLCRQ 940
Db 841 EDTCHSRREVGVVCSRYTDLVRLVNGKSCDGOVEINVLGHWSLCTDTHWDPEDARVLCRQ 900
QY 941 LSCGTALSTGGKYIGERSVRVWGHREHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICT 1000
Db 901 LSCGTALSTGGKYIGERSVRVWGHREHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICT 960
QY 1001 GSLTQPLFCLANVSDPYLSAVPEGSALICLEDKRLRLVGDSCRAGRVEIYHDFGWTI 1060
Db 961 GSLTQPLFCLANVSDPYLSAVPEGSALICLEDKRLRLVGDSCRAGRVEIYHDFGWTI 1020
QY 1061 CDGWDLSDAHVVYVQKLGCGVAFNATVSAHFEGGSGPIWLLDNLCTGTESHLWQCPSCRGW 1120
Db 1021 CDGWDLSDAHVVYVQKLGCGVAFNATVSAHFEGGSGPIWLLDNLCTGTESHLWQCPSCRGW 1080
QY 1121 GQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVYNGTWGSGVRRNITTAIGIV 1180
Db 1081 GQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVYNGTWGSGVRRNITTAIGIV 1140
QY 1181 CROLGCGENGVS LAPLSKTGSGFMWVDIQC PKTHISTWQCLSAFWERRISSPAETWI 1240
Db 1141 CROLGCGENGVS LAPLSKTGSGFMWVDIQC PKTHISTWQCLSAFWERRISSPAETWI 1200
QY 1241 TCEDRVRVGGDETCGRVEIWHAGSWGTVCDSDWDLAEAVVYVQKLGCGSALALRDAS 1300
Db 1201 TCEDRVRVGGDETCGRVEIWHAGSWGTVCDSDWDLAEAVVYVQKLGCGSALALRDAS 1260
QY 1301 FQGTGCTIWLDDMRCKGNESFLWDCHAKPWQSDCGKHEDAGVRCGGSLKSLNASSGHL 1360
Db 1261 FQGTGCTIWLDDMRCKGNESFLWDCHAKPWQSDCGKHEDAGVRCGGSLKSLNASSGHL 1320
QY 1361 ALILSSIFGLLILVLFILFTWCRVQKOKHPLRVSTRRGSLEENLFHEMETCLKREDP 1420

Db 1321 ALILSSIFGLLILVLFILFTWCRVQKOKHPLRVSTRRGSLEENLFHEMETCLKREDP 1380
QY 1421 HGTRTSDTPNHGCEADSDTSLGVLPAASEATK 1453
Db 1381 HGTRTSDTPNHGCEADSDTSLGVLPAASEATK 1413
RESULT 4
US-09-759-130B-383
; Sequence 383, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: MP100-5350NM1
; CURRENT APPLICATION NUMBER: US/09759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 383
; LENGTH: 1413
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-130B-383

Query Match 97.2%; Score 7913; DB 9; Length 1413;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 41 FNGTDLRLVNGDPCSGTVEVKFGQGWTVCDGWNNTASTVYVCKQLGCGFSFAMFRF 100
Db 1 FNGTDLRLVNGDPCSGTVEVKFGQGWTVCDGWNNTASTVYVCKQLGCGFSFAMFRF 60
QY 101 GOAVTRHGKIWLDDVSCYGNESALWECOHREWSHNCYHGEDVGYNCYGEANLRLVDG 160
Db 61 GOAVTRHGKIWLDDVSCYGNESALWECOHREWSHNCYHGEDVGYNCYGEANLRLVDG 120
QY 161 NNSCSGRVEVKFQERWGTICDDGWNLTAAVYVCRQLGCGFSFISGVVNSPAVLRPIWLD 220
Db 121 NNSCSGRVEVKFQERWGTICDDGWNLTAAVYVCRQLGCGFSFISGVVNSPAVLRPIWLD 180

QY 221 DILCOGNELALWNCRRHGWGNDHCSHNEEDVLTLCYDSDLELRLVGGTNRNCRMGVELKIQ 280
DB 181 DILCOGNELALWNCRRHGWGNDHCSHNEEDVLTLCYDSDLELRLVGGTNRNCRMGVELKIQ 240
QY 281 GRWGTVCCHHKNNAADAVCKQLGCGTALHPAGLPHLOSGSDVWLVLDGVSCEGNESEFLWD 340
DB 241 GRWGTVCCHHKNNAADAVCKQLGCGTALHPAGLPHLOSGSDVWLVLDGVSCEGNESEFLWD 300
QY 341 CRHSGTVNFDCLHONDVSVICSGADLELRLADGNNCSGRVEVRIHEQWNTICDQWKN 400
DB 301 CRHSGTVNFDCLHONDVSVICSGADLELRLADGNNCSGRVEVRIHEQWNTICDQWKN 360
QY 401 EQALVVCCKQLGCGPFSVFGSRRAPSPNEARDIWNISICTGNESALWDCYDGRKARTCFR 460
DB 361 EQALVVCCKQLGCGPFSVFGSRRAPSPNEARDIWNISICTGNESALWDCYDGRKARTCFR 420
QY 461 RSDAGVTCSDKADLDRLVGAHSPCYGRLEVYQGEWGTVCCHDRWSTRNAVVCCKQLGCG 520
DB 421 RSDAGVTCSDKADLDRLVGAHSPCYGRLEVYQGEWGTVCCHDRWSTRNAVVCCKQLGCG 480
QY 521 KPHVFGMTYFKASGPIWDDVSCIGNESNIWDCHEHSGKHCNVHREDVIVTCSGDAT 580
DB 481 KPHVFGMTYFKASGPIWDDVSCIGNESNIWDCHEHSGKHCNVHREDVIVTCSGDAT 540
QY 581 WGLRLVGGSNRCGRLEVYFQGRWGTVCDDGWNKAAAVVCSQLDCPSSIIIGMGLNAST 640
DB 541 WGLRLVGGSNRCGRLEVYFQGRWGTVCDDGWNKAAAVVCSQLDCPSSIIIGMGLNAST 600
QY 641 GYCKIWLDDVSCDGEDSLNCSRNGNNDCHSDSHEDVGVICSDASDMLRLVGGSSRCA 700
DB 601 GYCKIWLDDVSCDGEDSLNCSRNGNNDCHSDSHEDVGVICSDASDMLRLVGGSSRCA 660
QY 701 GKVEVNVQGAIVGILCANGWGMNTAEVVCROLECGSARVRSREPHFTERTLHILMSNGCT 760
DB 661 GKVEVNVQGAIVGILCANGWGMNTAEVVCROLECGSARVRSREPHFTERTLHILMSNGCT 720
QY 761 GGEASLWDCIRWEWKOTACHLNMEASLICSASRQPRLVGADMPGCSGRVEVKHADTWRSVC 820
DB 721 GGEASLWDCIRWEWKOTACHLNMEASLICSASRQPRLVGADMPGCSGRVEVKHADTWRSVC 780
QY 821 DSDFSLHAANVLCRELNGDAISLSVGDHFGKGNGLTWAERFQEGSETHALALCPVQHP 880
DB 781 DSDFSLHAANVLCRELNGDAISLSVGDHFGKGNGLTWAERFQEGSETHALALCPVQHP 840
QY 881 EDTCIHSREVGVCSTYDRLVNGKSCQDGOVEINVLGHWSLUCDTHWDPEDARVLRCQ 940
DB 841 EDTCIHSREVGVCSTYDRLVNGKSCQDGOVEINVLGHWSLUCDTHWDPEDARVLRCQ 900
QY 941 LSCGTALSTTGGKYIGRSVRVMGHRPHCLGNESLDDNCOMVTVLGAPPCIHGNTVSIVCT 1000
DB 901 LSCGTALSTTGGKYIGRSVRVMGHRPHCLGNESLDDNCOMVTVLGAPPCIHGNTVSIVCT 960
QY 1001 GSLTQPLFPCLANVSDPYLSAVPGESALICLEDKRLRLVDGDSRACGRVEIYHDFGWTI 1060
DB 961 GSLTQPLFPCLANVSDPYLSAVPGESALICLEDKRLRLVDGDSRACGRVEIYHDFGWTI 1020
QY 1061 CDDGWLDSDARHVVCKLGGCGVAFNATVSAHFEGSGPIWDDNCTGTESHLWQCPSRGW 1120
DB 1021 CDDGWLDSDARHVVCKLGGCGVAFNATVSAHFEGSGPIWDDNCTGTESHLWQCPSRGW 1080
QY 1121 GQHDCRHKEDAGVTCSEFTALRLYSETETESACGRLEVYNGTWGSGVRRNITTAIAGV 1180
DB 1081 GQHDCRHKEDAGVTCSEFTALRLYSETETESACGRLEVYNGTWGSGVRRNITTAIAGV 1140
QY 1181 CRQLGCCGNGVWSLAPLSKTSKTSRGMVWDDDIQCPKTHISIWQCLSAPHERLSSPAETWI 1240
DB 1141 CRQLGCCGNGVWSLAPLSKTSKTSRGMVWDDDIQCPKTHISIWQCLSAPHERLSSPAETWI 1200
QY 1241 TCEDRIRVRGDTFCGRVEIWHAGSNGTVCDDSDWDLAEAEVVCQOLGCCSALALRDAS 1300
DB 1201 TCEDRIRVRGDTFCGRVEIWHAGSNGTVCDDSDWDLAEAEVVCQOLGCCSALALRDAS 1260

QY 1301 FGOGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCHKEHDAGVRCGQSLKSNASSGHL 1360
DB 1261 FGOGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCHKEHDAGVRCGQSLKSNASSGHL 1320
QY 1361 ALILSSIFGLLLVFLFLTWCRVQKQKHLPLRVSTRRRGSLLENLFHEMETCLKREDP 1420
DB 1321 ALILSSIFGLLLVFLFLTWCRVQKQKHLPLRVSTRRRGSLLENLFHEMETCLKREDP 1380
QY 1421 HGRTSDTTPNHGCEADSDTSLIGVLPAASEATK 1453
DB 1381 HGRTSDTTPNHGCEADSDTSLIGVLPAASEATK 1413

RESULT 5

US-10-042-431-14
; Sequence 14, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: BRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-602
; CURRENT APPLICATION NUMBER: US/10/042,431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 14
; LENGTH: 1319
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-431-14

Query Match 91.3%; Score 7426; DB 9; Length 1319;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 41 FNGTDLRLVNGDPCSGTVEVKFQGGWGTVCDDGWNNTASTVVCCKQLGCGPFSFAMPRF 100
DB 1 FNGTDLRLVNGDPCSGTVEVKFQGGWGTVCDDGWNNTASTVVCCKQLGCGPFSFAMPRF 60
QY 101 GOATVRHGKIWLDDVSCYGNESALWECOHREWGSHNCHYHGEDVGVNCGYGEANLRLVYDG 160
DB 61 GOATVRHGKIWLDDVSCYGNESALWECOHREWGSHNCHYHGEDVGVNCGYGEANLRLVYDG 120
QY 161 NNSCSGRVEVKFQBRWGTICDDGWNLTAAVVCROLCGCPSSFISGGVNSPAVLRIWLD 220
DB 121 NNSCSGRVEVKFQBRWGTICDDGWNLTAAVVCROLCGCPSSFISGGVNSPAVLRIWLD 180
QY 221 DILCOGNELALWNCRRHGWGNDHCSHNEEDVLTLCYDSDLELRLVGGTNRNCRMGVELKIQ 280
DB 181 DILCOGNELALWNCRRHGWGNDHCSHNEEDVLTLCYDSDLELRLVGGTNRNCRMGVELKIQ 240
QY 281 GRWGTVCCHHKNNAADAVCKQLGCGTALHPAGLPHLOSGSDVWLVLDGVSCEGNESEFLWD 340
DB 241 GRWGTVCCHHKNNAADAVCKQLGCGTALHPAGLPHLOSGSDVWLVLDGVSCEGNESEFLWD 300
QY 341 CRHSGTVNFDCLHONDVSVICSGADLELRLADGNNCSGRVEVRIHEQWNTICDQWKN 400
DB 301 CRHSGTVNFDCLHONDVSVICSGADLELRLADGNNCSGRVEVRIHEQWNTICDQWKN 360
QY 401 EQALVVCCKQLGCGPFSVFGSRRAPSPNEARDIWNISICTGNESALWDCYDGRKARTCFR 460
DB 361 EQALVVCCKQLGCGPFSVFGSRRAPSPNEARDIWNISICTGNESALWDCYDGRKARTCFR 420
QY 461 RSDAGVTCSDKADLDRLVGAHSPCYGRLEVYQGEWGTVCCHDRWSTRNAVVCCKQLGCG 520

Db 421 RSDAGVICSADKADLDRLVLGHAHSPCYGRLEVKYQEGWGTCHDRWSTRNAAVVCKQLGCG 480
Qy 521 KPMHVFMTYKEASGPIWLDVSCIGNESINWDEHSGWGHKNCVHREDVIVTCSGDAT 580
Db 481 KPMHVFMTYKEASGPIWLDVSCIGNESINWDEHSGWGHKNCVHREDVIVTCSGDAT 540
Qy 581 WGLRLVGSNRCSGRLEVYFQGRWGTVCDDGWNKAAAVVCSQDSCPSIITGMGLGNAST 640
Db 541 WGLRLVGSNRCSGRLEVYFQGRWGTVCDDGWNKAAAVVCSQDSCPSIITGMGLGNAST 600
Qy 641 GYGKTLDDVSCDGDSDSLWSCRNSGWNDCSHSDVGVICSDASDMLRLVGSRSRA 700
Db 601 GYGKTLDDVSCDGDSDSLWSCRNSGWNDCSHSDVGVICSDASDMLRLVGSRSRA 660
Qy 701 GKVEVNOGAVGILCANGWGNIAEVVCRQLECGSAIRVSREPHTERTLHLMSNSCT 760
Db 661 GKVEVNOGAVGILCANGWGNIAEVVCRQLECGSAIRVSREPHTERTLHLMSNSCT 720
Qy 761 GGEASLWPCIRWEMKOTACHLNMEASLCSAHRQPRLVGADMPGSGRVEVVKHDTWRSVC 820
Db 721 GGEASLWPCIRWEMKOTACHLNMEASLCSAHRQPRLVGADMPGSGRVEVVKHDTWRSVC 780
Qy 821 DSDFSLHAANVLCRELNGDALSISVGDHFGKNGLTWAEKFCQBSGSETHALCPVQHP 880
Db 781 DSDFSLHAANVLCRELNGDALSISVGDHFGKNGLTWAEKFCQBSGSETHALCPVQHP 840
Qy 881 EDTCHSREVGVVCSRYTDVRLVNGKSCDGOVEINVLGHWSGLCDTHWDPEDARVLCRQ 940
Db 841 EDTCHSREVGVVCSRYTDVRLVNGKSCDGOVEINVLGHWSGLCDTHWDPEDARVLCRQ 900
Qy 941 LSCGTALSTTGKYGISRSVRVWGHRRFCLGNESLDDNCOMTVLGAPCIHGNTVSVICT 1000
Db 901 LSCGTALSTTGKYGISRSVRVWGHRRFCLGNESLDDNCOMTVLGAPCIHGNTVSVICT 960
Qy 1001 GSLTOPLPCLANVSDPYLSAYPEGSALICLEDKRLRLVDGDSRCAGRVEVYHDFGWTI 1060
Db 961 GSLTOPLPCLANVSDPYLSAYPEGSALICLEDKRLRLVDGDSRCAGRVEVYHDFGWTI 1020
Qy 1061 CDGWDLSDAHVVCKLGGVAFNATVSAHFEGSGPIWLDLNCGTGESHLMQCPSRGW 1120
Db 1021 CDGWDLSDAHVVCKLGGVAFNATVSAHFEGSGPIWLDLNCGTGESHLMQCPSRGW 1080
Qy 1121 GOHDCRHKEDAGVICSSETALRLYSETETESCAGRLEVYFNGTWSGRRNITTAIGIV 1180
Db 1081 GOHDCRHKEDAGVICSSETALRLYSETETESCAGRLEVYFNGTWSGRRNITTAIGIV 1140
Qy 1181 CRQLGCGENGVSAPLSKTSKSGFMWVDDIQCPKTHISIWOCLSAWERRISSPAEETWI 1240
Db 1141 CRQLGCGENGVSAPLSKTSKSGFMWVDDIQCPKTHISIWOCLSAWERRISSPAEETWI 1200
Qy 1241 TCEDRIRVRGGTSGRVEIWHAGSWGTVCDSDWDLAEAEVVCQOLGCGSALAALRAS 1300
Db 1201 TCEDRIRVRGGTSGRVEIWHAGSWGTVCDSDWDLAEAEVVCQOLGCGSALAALRAS 1260
Qy 1301 FCQGTGTWLDMDRCKGNESFLWDCCHAKPWGSDCGHREDAGVRCGSLKSLNASSGH 1359
Db 1261 FCQGTGTWLDMDRCKGNESFLWDCCHAKPWGSDCGHREDAGVRCGSLKSLNASSGH 1319

RESULT 6

US-09-759-130b-384

; Sequence 384, Application US/09759130B

; Publication No. US2003002279A1

; GENERAL INFORMATION:

; APPLICANT: Millennium Pharmaceuticals, Inc.

; APPLICANT: McCarthy, Sean A

; APPLICANT: Fraser, Christopher C

; APPLICANT: Sharp, John D

; APPLICANT: Barnes, Thomas S

; APPLICANT: Kirt, Susan J

; APPLICANT: Mackay, Charles R

; APPLICANT: Myers, Paul S

; APPLICANT: Leiby, Kevin R

; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: MF100-5350MIM
; CURRENT APPLICATION NUMBER: US/09759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 384
; LENGTH: 1319
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-130b-384

Query Match 91.3%; Score 7426; DB 9; Length 1319;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 FNGTDLRLVNGDPCSGTVEVKFQGGWGTVCDDGWNNTASTVVCCKQLGCPFSFAMPRF 100
Db 1 FNGTDLRLVNGDPCSGTVEVKFQGGWGTVCDDGWNNTASTVVCCKQLGCPFSFAMPRF 60
Qy 101 GOAVTRHGKIWLDDVSCYGNESALWECQHREWSHNCYHGEDVGVNCGEANLGLRLVDG 160
Db 61 GOAVTRHGKIWLDDVSCYGNESALWECQHREWSHNCYHGEDVGVNCGEANLGLRLVDG 120
Qy 161 NNCSGRVEVKFOERWGTICDDGWNLTAAVVCROLGCPSSFISSGVVNSPAVLRPIWLD 220
Db 121 NNCSGRVEVKFOERWGTICDDGWNLTAAVVCROLGCPSSFISSGVVNSPAVLRPIWLD 180
Qy 221 DILCOGNELALWNCRRHGWGNDHCSHNEEDVTITCYDSSDLELRLVGGTNRCGRVELKIQ 280
Db 181 DILCOGNELALWNCRRHGWGNDHCSHNEEDVTITCYDSSDLELRLVGGTNRCGRVELKIQ 240
Qy 281 GRWGTVCCHKNNAAADVVCCKQLGCGTALHFAGLPHLQSGSDVWLDGVSCSGNESFLWD 340
Db 241 GRWGTVCCHKNNAAADVVCCKQLGCGTALHFAGLPHLQSGSDVWLDGVSCSGNESFLWD 300
Qy 341 CRHSGTVNFDCLHONDVSVICSDGADLELRLADGNNCSGRVEVRHIEQWWTICDQNNKN 400
Db 301 CRHSGTVNFDCLHONDVSVICSDGADLELRLADGNNCSGRVEVRHIEQWWTICDQNNKN 360
Qy 401 EQALVVCCKQLGCPFSVFGSRRRAKPSNEARDIWINISICTGNESALWDCDCTYDCAKRTCFR 460
Db 361 EQALVVCCKQLGCPFSVFGSRRRAKPSNEARDIWINISICTGNESALWDCDCTYDCAKRTCFR 420
Qy 461 RSDAGVICSADKADLDRLVLGHAHSPCYGRLEVKYQEGWGTCHDRWSTRNAAVVCKQLGCG 520
Db 421 RSDAGVICSADKADLDRLVLGHAHSPCYGRLEVKYQEGWGTCHDRWSTRNAAVVCKQLGCG 480

Qy 521 KPMHVFMTYKEASGPIWLDVSCIGNESINWDCSHGSGKHCNVHREDVIVTCSDGAT 580
Db 481 KPMHVFMTYKEASGPIWLDVSCIGNESINWDCSHGSGKHCNVHREDVIVTCSDGAT 540
Qy 581 WGLRLVGSNRCGRLEVYFQGRNGTVCDGWSKAAAVVCSQDPCSSIIIGMGLNAST 640
Db 541 WGLRLVGSNRCGRLEVYFQGRNGTVCDGWSKAAAVVCSQDPCSSIIIGMGLNAST 600
Qy 641 GYGIWLDVSDGDESDMSCRSGNNDCHSDSHEDVGVICSDASDMELRLVGSRRCA 700
Db 601 GYGIWLDVSDGDESDMSCRSGNNDCHSDSHEDVGVICSDASDMELRLVGSRRCA 660
Qy 701 GKVEVNVQAGVILCANGWGNIAEVVCRQLCEGSAIRVSREPHTERTLHILMSNGCT 760
Db 661 GKVEVNVQAGVILCANGWGNIAEVVCRQLCEGSAIRVSREPHTERTLHILMSNGCT 720
Qy 761 GGEASLWDCIRWEKOTACHLNEASLICSARHQPRLVGADMPGSRVEVKAHTWRSVC 820
Db 721 GGEASLWDCIRWEKOTACHLNEASLICSARHQPRLVGADMPGSRVEVKAHTWRSVC 780
Qy 821 DSDFSLHAANYLCLRELNCGDAISLSVGDHFGKGNLTWAEKFCQEGSETHLALCPVQHP 880
Db 781 DSDFSLHAANYLCLRELNCGDAISLSVGDHFGKGNLTWAEKFCQEGSETHLALCPVQHP 840
Qy 881 EDTCHSREVGVVCSRYTDLRLVNGKSQCDCQVEINVLGHWSGLCDTHWDPEDARVLCRQ 940
Db 841 EDTCHSREVGVVCSRYTDLRLVNGKSQCDCQVEINVLGHWSGLCDTHWDPEDARVLCRQ 900
Qy 941 LSCGTALSTTGGKYIGERSVRWGHFRHCLGNESLDDNCQMTVLGAPPCIHGNTVSIVCT 1000
Db 901 LSCGTALSTTGGKYIGERSVRWGHFRHCLGNESLDDNCQMTVLGAPPCIHGNTVSIVCT 960
Qy 1001 GSLTQPLFPCLANVSDPYSAPVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDFGWTG 1060
Db 961 GSLTQPLFPCLANVSDPYSAPVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDFGWTG 1020
Qy 1061 CDDGWDLSDAHVCOKLGCQVAFNATVSAHFGEGSGPIWLDLNCGTGTFESHLMWCPSRGW 1120
Db 1021 CDDGWDLSDAHVCOKLGCQVAFNATVSAHFGEGSGPIWLDLNCGTGTFESHLMWCPSRGW 1080
Qy 1121 GOHDCRHKEDAGVICSEPTALRLYSETETESCAGRLEVFYNGTWGSGVGRNITTAIAGIV 1180
Db 1081 GOHDCRHKEDAGVICSEPTALRLYSETETESCAGRLEVFYNGTWGSGVGRNITTAIAGIV 1140
Qy 1181 CRQLCGGNGVYSLAPLSKTSKSGFWVDDIOCPKTHISWOCLSAPWERRISSPAETWI 1240
Db 1141 CRQLCGGNGVYSLAPLSKTSKSGFWVDDIOCPKTHISWOCLSAPWERRISSPAETWI 1200
Qy 1241 TCEDRIRVRGGDTESGSRVEIWHAGSWGTVCDSDSLAEAEVVCQQLCGSALAALRDAS 1300
Db 1201 TCEDRIRVRGGDTESGSRVEIWHAGSWGTVCDSDSLAEAEVVCQQLCGSALAALRDAS 1260
Qy 1301 FGQGTGTTWLDMMCKGNESFLWCHAKPWGQSCGCHKEDAGVRCSGOSLSKLNASSGH 1359
Db 1261 FGQGTGTTWLDMMCKGNESFLWCHAKPWGQSCGCHKEDAGVRCSGOSLSKLNASSGH 1319

RESULT 7

US-09-977-577-13
; Sequence 13, Application US/09977577
; Patent No. US2002015595A1
; GENERAL INFORMATION:
; APPLICANT: MOESTRUP, Soren
; APPLICANT: MOLLER, Holger J.
; TITLE OF INVENTION: THE FUNCTION OF A HAPTOGLOBIN-HAEMOGLOBIN RECEPTOR AND THE USES THEREOF
; FILE REFERENCE: MOESTRUP-1A
; CURRENT APPLICATION NUMBER: US/09/977.577
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: US 60/270,120
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: DK PA 2001 00039
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: DK PA 2000 01543

; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 1151
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-977-577-13

Query Match 42.3%; Score 3441.5; DB 9; Length 1151;
Best Local Similarity 55.2%; Pred. No. 9e-239;
Matches 608; Conservative 166; Mismatches 311; Indels 17; Gaps 6;

Qy 364 GADLELRADSNCSGRVEVRIHDEQMTICDQNNKNEQALVYVQKQCPESVFGSRRAK 423
Db 41 GADLELRADSNCSGRVEVRIHDEQMTICDQNNKNEQALVYVQKQCPESVFGSRRAK 100
Qy 424 PSNEARDIWNISICTGNESALWDCYDYGKAKRT-CFRRSDAGVTCSDKADLDLRLVGAH 482
Db 101 SSAGSGRIWMDHVSRCGNESALWDCYDYGKAKRT-CFRRSDAGVTCSDKADLDLRLVGAH 160
Qy 483 SPCYGRLEVKYQGGMTVCHDRWSTRNAAVVCKQLCGCKPMPHVFMTYFKEASGPIWLD 542
Db 161 SPCYGRLEVKYQGGMTVCHDRWSTRNAAVVCKQLCGCKPMPHVFMTYFKEASGPIWLD 220
Qy 543 VSCIGNESINWDCSHGSGKHCNVHREDVIVTCSDGATMGLRLVGSNRCGRLEVYFQ 602
Db 221 VSCIGNESINWDCSHGSGKHCNVHREDVIVTCSDGATMGLRLVGSNRCGRLEVYFQ 280
Qy 603 RWGTVCDGWSKAAAVVCSQDPCSSIIIGMGLNASTGYGKIWLDVSDGDESDLWASC 662
Db 281 RWGTVCDGWSKAAAVVCSQDPCSSIIIGMGLNASTGYGKIWLDVSDGDESDLWASC 340
Qy 663 ENSGNGNDCSHSDGVVCSQDPCSSIIIGMGLNASTGYGKIWLDVSDGDESDLWASC 722
Db 341 ENSGNGNDCSHSDGVVCSQDPCSSIIIGMGLNASTGYGKIWLDVSDGDESDLWASC 400
Qy 723 LAEVVCRQLCGSALTSVQVYKIQATNTWFLSSCNGNETSLWDCRKNWQGLTCDHY 782
Db 401 LAEVVCRQLCGSALTSVQVYKIQATNTWFLSSCNGNETSLWDCRKNWQGLTCDHY 460
Qy 783 NEASLICSARHQPRLVGADMPGSRVEVKAHTWRSVCDSDSLHAANYLCLRELNCGDAI 842
Db 461 NEASLICSARHQPRLVGADMPGSRVEVKAHTWRSVCDSDSLHAANYLCLRELNCGDAI 520
Qy 843 SLSVGDHFGKGNLTWAEKFCQEGSETHLALCPVQHPEDTCHSREVGVVCSRYTDLRL 902
Db 521 SLSVGDHFGKGNLTWAEKFCQEGSETHLALCPVQHPEDTCHSREVGVVCSRYTDLRL 580
Qy 903 YNGKSCQDQVBEINVLGHWSGLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVR 962
Db 581 YNGKSCQDQVBEINVLGHWSGLCDTHWDPEDARVLCRQLSCGTALSTTGGKYIGERSVR 640
Qy 963 WGRHFCILGNESLDDNCQMTVLGAPPCIHGNTVSIVCTGSLTQPLFPCLANVSDPYSAP 1022
Db 641 WGRHFCILGNESLDDNCQMTVLGAPPCIHGNTVSIVCTGSLTQPLFPCLANVSDPYSAP 700
Qy 1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDFGWTICDDGWDLSDAHVCOKLGCQV 1082
Db 701 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDFGWTICDDGWDLSDAHVCOKLGCQV 760
Qy 1083 FNATVSAHFGEGSGPIWLDLNCGTGTFESHLMWCPSRGVGRNITTAIAGIV 1142
Db 761 FNATVSAHFGEGSGPIWLDLNCGTGTFESHLMWCPSRGVGRNITTAIAGIV 820
Qy 1143 IYSETETESCAGRLEVFYNGTWGSGVGRNITTAIAGIV 1202
Db 821 IYSETETESCAGRLEVFYNGTWGSGVGRNITTAIAGIV 880
Qy 1203 GPMWVDDIOCPKTHISWOCLSAPWERRISSPAETWITCEDRIRVRGGDTESGSRVEI 1262
Db 881 GPMWVDDIOCPKTHISWOCLSAPWERRISSPAETWITCEDRIRVRGGDTESGSRVEI 940

QY 1263 HAGSWGTCDDSDWDLAEAEVWCOQLGCGSALAAALRDASFGOGGTGTWLDMDRCCKNESFL 1322
DB 941 HGGSWGTCDDSDWDLDDAQVVCQQLGCGPALKAFKAEAFEGOGGTGTWLDNEVKCKGNESL 1000
QY 1323 WDCHAKPWGOSCGHEDAGVRCQSGLK-----SLNASSGHLALILSSIFGLLLVLF 1376
DB 1001 WDCPARRWGHSECGHEDAAVNCCTDISVQKTPQKATGRSSQSSFIAGVILGVVLLAIF 1060
QY 1377 I--LFTWCRVOKKHLPLRVSTRRRGSLLENLFHEMETCKLREDPHGTRTSDDT--PNHGC 1434
DB 1061 VALFELTKRRRQR---LAVSSRGENLVHQIYQREMNCLNADDLDMNSEN--SHES 1115
QY 1435 EDASDTSLGV---LPASATK 1453
DB 1116 ADFSAAELISVSKFLPISGMK 1137

RESULT 8
US-09-977-577-10
; Sequence 10, Application US/09977577
; Patent No. US20020155995A1
; GENERAL INFORMATION:
; APPLICANT: MOESTRUP, Soren
; APPLICANT: MOLLER, Holger J.
; TITLE OF INVENTION: THE FUNCTION OF A HAPTOGLOBIN-HAEMOGLOBIN RECEPTOR AND THE USES
; FILE REFERENCE: MOESTRUP-1A
; CURRENT APPLICATION NUMBER: US/09/977,577
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: US 60/270,120
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: DK PA 2001 00039
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: DK PA 2000 01543
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 1116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-977-577-10

Query Match 42.1%; Score 3429; DB 9; Length 1116;
Best Local Similarity 56.2%; Pred. No. 6.8e-238;
Matches 598; Conservative 161; Mismatches 294; Indels 12; Gaps 4;

QY 364 GADLELRADGNNCGSRVEVRIHEOWTTCIDQNNKNEQALVVCQKOLGCPFSVFGSRRAK 423
DB 41 GTDKELRLVDGNNCKSGRVEVKVQEWGTVCNNGWSMEAVSVICNLQGCPTAIKAPGWAN 100
QY 424 PSNEARDIWNISCTGNESALWDCTYDGAART-CFRSDAGVICSADKADLRLVGAH 482
DB 101 SSAGSGRIWMDHVSRCGNESALWDCKDHWGKHSNTHQDAGVTCSDGNSLMLRTRGG 160
QY 483 SPYGLREVKYOGWGTVDHNRSTRNAAVVCQKOLGCGPMMVFMGTYPKEASGPTWLD 542
DB 161 NMCSGRIEIKFQWGTVDHNRSTRNAAVVCQKOLGCGPMMVFMGTYPKEASGPTWLD 220
QY 543 VSCIGNESINWDCHEHSGWCKHNCVHREDVITVCSGDTWGLRLVGGSNRCGRLEVYFQ 602
DB 221 LINCNGESALWCKHOGWCKHNCVHREDVITVCSGDTWGLRLVGGSNRCGRLEVYFQ 280
QY 603 RMGTVCDDGWNKAAAVVCQKOLGCGPMMVFMGTYPKEASGPTWLD 562
DB 281 EMGTICDDGWDSDAACKVQKGLCPATVTAIGRVNASKGFGHVLWDSVSCQHEPAAWOC 340
QY 663 RNSGWNDCSHEDVGVTCSDASDMLRLVGGSNRCGRLEVYFQ 722
DB 341 KHHEWGHKYNHEDAGVTCSDGSLRLVGGSNRCGRLEVYFQ 400
QY 723 IAEVVCQKOLGCGSALATYQVYSKIQTATWFLSSCNGNETSLMDCKKNWQGLTCDHY 782
DB 401 EADVVCQKOLGCGSALATYQVYSKIQTATWFLSSCNGNETSLMDCKKNWQGLTCDHY 460

QY 783 MEASLICSARHQPRLVGADMPGCSGRVEVKHADTWRSVCDSDFSLHAANYLCRELNGDAI 842
DB 461 BEAKITCSARHQPRLVGADMPGCSGRVEVKHADTWRSVCDSDFSLHAANYLCRELNGDAI 520
QY 843 SLVSVDHKGKNGLTWAEKFCQEGSETHALCPIVQHPEDTCHHSREVGVCVSRITDVL 902
DB 521 SLTGAHFGEGNGQIWAEEFQCEHSHLSLCPVAPRPEGTCHSHRDVGVCSRYTEIL 580
QY 903 VNGKSOCDQOEVINVLGHWGSLCDTHWDPEDARVLCRQLSCGTALSTGTGKVIERSVRV 962
DB 581 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQKLCGVALSTPGARFGKNGQI 640
QY 963 WGRFHCLGNESLNDNCOMTVLGAAPPCIHGNTVITCTSLTQPLFPCLANVSDPVLAV 1022
DB 641 WRHMFCTGTQOHMGDCPVTALGASLCPSEQVAVSCNQSQTLSNCSNSSLGPTPTI 700
QY 1023 PEGSALICLEDKRLVDCDSRCAGRVETIYHDGFWGTCIDCDGWDLSDAHVVCQKLCGVA 1082
DB 701 PEESAVACIESQQLRLVNGGRCAGRVETIYHGGWGTICDDSDWLDSDAHVVCQKLCGVA 760
QY 1083 FNATVSAHFEGSGPIWLDLNLCTGTESHLWQPCSRGWGQHCRCRKEDAGVICSEFTALR 1142
DB 761 INATGSAHFEGSGTPIWLDLNLCTGTESHLWQPCSRGWGQHCRCRKEDAGVICSEFTALR 820
QY 1143 LYSETETESACGRLEVYNGTWGVRNRNITTAIGVCRQLGCGGNGVVSAPLSKTS 1202
DB 821 LTSEAAREACGRLEVYNGTWGVRNRNITTAIGVCRQLGCGGNGVVSAPLSKTS 880
QY 1203 GFWMVDDIOTCPKTHISWCLSPAPWERRISSPAEETWITCEDRIRVGGDTSCSGRVEIW 1262
DB 881 IPMVDNVQCPKGPDTLWQCPSPWPKRLASPEETWITCEDRIRVGGDTSCSGRVEIW 940
QY 1263 HAGSWGTCDDSDWDLAEAEVWCOQLGCGSALAAALRDASFGOGGTGTWLDMDRCCKNESFL 1322
DB 941 HGGSWGTCDDSDWDLDDAQVVCQQLGCGPALKAFKAEAFEGOGGTGTWLDNEVKCKGNESL 1000
QY 1323 WDCHAKPWGOSCGHEDAGVRCQSGLK-----SLNASSGHLALILSSIFGLLLVLF 1376
DB 1001 WDCPARRWGHSECGHEDAAVNCCTDISVQKTPQKATGRSSQSSFIAGVILGVVLLAIF 1060
QY 1377 I--LFTWCRVOKKHLPLRVSTRRRGSLLENLFHEMETCKLRED 1419
DB 1061 VALFELTKRRRQR---LAVSSRGENLVHQIYQREMNCLNADD 1102

RESULT 9
US-09-977-577-12
; Sequence 12, Application US/09977577
; Patent No. US20020155995A1
; GENERAL INFORMATION:
; APPLICANT: MOESTRUP, Soren
; APPLICANT: MOLLER, Holger J.
; TITLE OF INVENTION: THE FUNCTION OF A HAPTOGLOBIN-HAEMOGLOBIN RECEPTOR AND THE USE
; FILE REFERENCE: MOESTRUP-1A
; CURRENT APPLICATION NUMBER: US/09/977,577
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: US 60/270,120
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: DK PA 2001 00039
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: DK PA 2000 01543
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 1156
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-977-577-12

Query Match 42.1%; Score 3429; DB 9; Length 1156;
Best Local Similarity 56.2%; Pred. No. 7.1e-238;

Matches 598; Conservative 161; Mismatches 294; Indels 12; Gaps 4;

Qy 364 GADLELRADGNNCGSRVEVRIHQWWTICDQNKNEQALVVKQKQCPFSVFGSRRAK 423
Db 41 GTDKELRLVDGENKSGRVEVKVQEWGTVCCNNGNSMEAVSVICNQLGCTPAIKAPGWAN 100
Qy 424 PSNEARDIWINISICTGNESALWDCYDYGKAKRT-CFRRSAGVYICSDKADLDLRLVGAH 482
Db 101 SSAGSGRIWMDHVSCRGNESALWCKHDGKHKHNSCTHQDAGVTCSDGNSLEMLRTRGG 160
Qy 483 SPCYGRLEVKYQWGTVCHDRWSTRNAVVKQKQCGKPMHVFMTYKEASGPIWLD 542
Db 161 NMCSGRIETKFGRRGTVCDDNFNDHASVICRQLECGSAVSFGSSNFGEGSGPIWFD 220
Qy 543 VSCIGNESINWDCSHSGKHKNCVHREDIVITCSGDATWGLRVGSGNRCSRLLEVYFQG 602
Db 221 LICNGESALWCKHOGKHKNCDAEDAGVICSKGADLSRLVDGVTECSRLEVRFG 280
Qy 603 RMTGVCDDGWSKAAAVVCSQIDCPSSITGMGLGNASTGYKGIWLDVSCDGDSDLSWC 662
Db 281 EMGTICDDGWSYDAVACKQKQCGPTAVTAIGRVNASKGFGHIWLDVSCOGHEPAVWQC 340
Qy 663 RNSGNGNDCSHSDVGYICSDASDMELRVGSSRCAGKVEVNVGAVGILCANGWGN 722
Db 341 KHHEWGHKHYCNHEDAGVTCSDGSDLELRGGGSRCACTVEVEIQRLLGKVCDRGWLK 400
Qy 723 TAEVVCRLQECGSALRVSRPHFTERTLHILMSNGCTGGESALWDCIRWENKQATACHLN 782
Db 401 EADVVCRLQCGSALKTSYQVYSKIQTNTWLFSSCNGNETSLMCKNWMGGTCDHY 460
Qy 783 MEASLICSNAHROPRLVGADMPGSRVEVKHADTWRSVCDSDFLHAAVYLCRELNCDAI 842
Db 461 BEAKITCSAHPREPRVLGGDIPCSGRVEVKHGTWGSICDSFSLAASVLCRELQGTIV 520
Qy 843 SLSVGDHFGKGNLTWAEKFCQEGSETHALCPIVQHPEPTCIHRSREVGVCSTYDRL 902
Db 521 STLGAHFEFEGNGQIWAEEFQEGHESHLSCPAPRPEGTCSHSDRVGVVCSRYTEIRL 580
Qy 903 VNGKSOCQGOVEINVLGHWSICDHPWDPEDARVLCRLSCGTALSTTGGKYIGERSVRV 962
Db 581 VNGKTPCGRVELKTLGAWGSLCNSHWDIEDHVLQQLKCGVALSTPGGARFGNGOI 640
Qy 963 WGRHFCILGNESLNDQNTVLGAPPCIHGNTVSVICTGSLTQPLFPCPLANVSDPYLSAV 1022
Db 641 WRHMFCTGTQEHMGDCPTALGASLCPSEQVASVICSNGQSOTLSSCNSSSLGPTPTI 700
Qy 1023 PEGSALICLEDKRLRLVDGSDRCAGRVYIHDGFWGTICDDGWDLSDAHVVCQKLGCGVA 1082
Db 701 PEESAVACIESQLRLVNGGRCAGRVYIHBGWSGTICDDSDWLDSDAHVVCRLGCGEA 760
Qy 1083 FNATVSAHFGGSGPIWLDLNCCTESHLMQCPSPRGWQHDCRHKEDAGVICSFTALR 1142
Db 761 INATCSAHFEGGTGPIWLDKCKNKESRIWQCHSHGQOQNCRIKEDAGVICSFMSLR 820
Qy 1143 LYSETETESACRLEVFYNGWGSVGRNITTAIAGIVCRLQCGGNGVVS LAPLSKGS 1202
Db 821 LTSEASREACAGRLEVFYNGWGTGKSSMETTVGVVVCRLQCGADKGINPASLDKMS 880
Qy 1203 GFMMVDDIQCPTHTSIWQCLSAHPERLSSPAETWITCEDRIIRVRGDTCECSRVEIW 1262
Db 881 IPMVNDVQCCKPGPTLWQCPSPSPWEKRLASPEETWITCONKIRLQEGPTSCSRVEIW 940
Qy 1263 HAGSGTVCDDSDWDLAEAEVVCQQLGCGSALAALRADSFQGTGTIWLDDMRCKKNESFL 1322
Db 941 HGGSGTVCDDSDWDLDDAQVVCQQLGCGPALKAPAEAFEGGTGPIWLNVEVKCKNESL 1000
Qy 1323 WDCHAKPWQDCGCHKEDAGVRCQSGLK-----SLNASSGLALILSSIFGLLLVLF 1376
Db 1001 WDCPARRGHSGCHGKEDAAVNCNTDISVQKTPQKATTCGRSSRQSSFFIAVGLGVLLAIF 1060
Qy 1377 I--LFLTWCVRQVKQKHLPLRVSTRRRGSLLENLFHEMETCLKRED 1419
Db 1061 VALFELTKRRQROR---LAVSSRGENLVHQIYREMNNSCLNADD 1102

RESULT 10

US-09-977-577-11
; Sequence 11, Application us/09977577
; Patent No. US2002015995A1
; GENERAL INFORMATION:
; APPLICANT: MOESTRUP, Soren
; APPLICANT: MOLLER, Holger J.
; TITLE OF INVENTION: THE FUNCTION OF A HAPTOGLOBIN-HAEMOGLOBIN RECEPTOR AND THE USE
; FILE REFERENCE: MOESTRUP-1A
; CURRENT APPLICATION NUMBER: US/09/977,577
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: US 60/270,120
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: DK PA 2001 00039
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: DK PA 2000 01543
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 1149
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-977-577-11

Query Match 41.7%; Score 3396.5; DB 9; Length 1149;
Best Local Similarity 54.4%; Pred. No. 1.5e-235;
Matches 597; Conservative 161; Mismatches 295; Indels 45; Gaps 5;

Qy 364 GADLELRADGNNCGSRVEVRIHQWWTICDQNKNEQALVVKQKQCPFSVFGSRRAK 423
Db 41 GTDKELRLVDGENKSGRVEVKVQEWGTVCCNNGNSMEAVSVICNQLGCTPAIKAPGWAN 100
Qy 424 PSNEARDIWINISICTGNESALWDCYDYGKAKRT-CFRRSAGVYICSDKADLDLRLVGAH 482
Db 101 SSAGSGRIWMDHVSCRGNESALWCKHDGKHKHNSCTHQDAGVTCSDGNSLEMLRTRGG 160
Qy 483 SPCYGRLEVKYQWGTVCHDRWSTRNAVVKQKQCGKPMHVFMTYKEASGPIWLD 542
Db 161 NMCSGRIETKFGRRGTVCDDNFNDHASVICRQLECGSAVSFGSSNFGEGSGPIWFD 220
Qy 543 VSCIGNESINWDCSHSGKHKNCVHREDIVITCSGDATWGLRVGSGNRCSRLLEVYFQG 602
Db 221 LICNGESALWCKHOGKHKNCDAEDAGVICSKGADLSRLVDGVTECSRLEVRFG 280
Qy 603 RMTGVCDDGWSKAAAVVCSQIDCPSSITGMGLGNASTGYKGIWLDVSCDGDSDLSWC 662
Db 281 EMGTICDDGWSYDAVACKQKQCGPTAVTAIGRVNASKGFGHIWLDVSCOGHEPAVWQC 340
Qy 663 RNSGNGNDCSHSDVGYICSDASDMELRVGSSRCAGKVEVNVGAVGILCANGWGN 722
Db 341 KHHEWGHKHYCNHEDAGVTCSDGSDLELRGGGSRCACTVEVEIQRLLGKVCDRGWLK 400
Qy 723 TAEVVCRLQECGSALRVSRPHFTERTLHILMSNGCTGGESALWDCIRWENKQATACHLN 782
Db 401 EADVVCRLQCGSALKTSYQVYSKIQTNTWLFSSCNGNETSLMCKNWMGGTCDHY 460
Qy 783 MEASLICSNAHROPRLVGADMPGSRVEVKHADTWRSVCDSDFLHAAVYLCRELNCDAI 842
Db 461 BEAKITCSAHPREPRVLGGDIPCSGRVEVKHGTWGSICDSFSLAASVLCRELQGTIV 520
Qy 843 SLSVGDHFGKGNLTWAEKFCQEGSETHALCPIVQHPEPTCIHRSREVGVCSTYDRL 902
Db 521 STLGAHFEFEGNGQIWAEEFQEGHESHLSCPAPRPEGTCSHSDRVGVVCSRYTEIRL 580
Qy 897 -----YDVRLVNKGKQCCQGOVEINVLGHWSICDTHW 929
Db 581 LIGSVTVKGTGLGSHSCLFLKPCLLPGYTEIRLVNKGKTPCEGRVELKTLGANGSLCNSHW 640
Qy 930 DPEDARVLCRLQCGSALTSTTGGKYIGERSVRWGRHFCILGNESLNDQNTVLGAPPC 989
Db 930 DPEDARVLCRLQCGSALTSTTGGKYIGERSVRWGRHFCILGNESLNDQNTVLGAPPC 989

Db 641 DIEDAHLVLCQKCGVALSTPCGAREFGKNGQWIRHMFHCTCTEHOHGDPCPTALGASLC 700
QY 990 IHGNTVSVICTGSLTOPLPFCIANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRAGRV 1049
Db 701 PSEQVASVTCGNSQSLSSNSSSLGPRTPIPESAVACTESQGLRLVNGGRCAGRV 760
QY 1050 EYHDFGWTICDDGWDLSDAHVCQKLCGVAFAFATVSAHFEGSGPTWLDLACTGTE 1109
Db 761 EYHESGWTICDDGWDLSDAHVCQKLCGVAFAFATVSAHFEGSGPTWLDLACTGTE 820
QY 1110 SHLWQPSRGWQHCHRHKEDAGVTCSEPTALRLYSETETESACRLEVFYNGTWSVGR 1169
Db 821 SRIWOCHSHGWQONCRHREDAVVICSEPMSLRLTSEASREACAGRLEVFYNGWGTGK 880
QY 1170 RNITTAIGIVCRQLCGGNGVVSAPLSKTCGPMWDDIOCPKTHISWQCLSPWHER 1229
Db 881 SSMSETVGVCRQLCGADKINPASPDLKAMSPMWDVQCPKGPDTLWQCPSPWPK 940
QY 1230 RISSPAEETWITCEDIRVRGGTSCSGRVEIWHAGSWGTCVDDSDMLAEAEVVCQOLGC 1289
Db 941 RLASPEETWITCDKIRLQEGTSCSGRVEIWHAGSWGTCVDDSDMLAEAEVVCQOLGC 1000
QY 1290 GSALALRASQGTGTWLDLMDCKGNEFLWDCHAKPWGSDCGHKEDAGVRCSGOS 1349
Db 1001 GPALKAFKEAFGGTGPWLNEVKCKGNESSLWDCPARRWGHSECGHKEDAAVNCCTDIS 1060
QY 1350 LK-----SLNASSGHLALISIFGLLLLVFI--LFTWCRVQKQKHLPLRVSTRRRG 1401
Db 1061 VQTPKATGTRSSROSSFAVGLGVLLAIFVALFLLTKRRQROR--LAVSSRGEN 1117
QY 1402 SLEENLFHEWETCLKRED 1419
Db 1118 LVHQIQYREMNNSCLNADD 1135

RESULT 11

; Sequence 78, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-602
; CURRENT APPLICATION NUMBER: US/10/042,431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 78
; LENGTH: 1436
; TYPE: PRT
; ORGANISM: Bos sp.
US-10-042-431-78

Query Match 37.9%; Score 3082; DB 9; Length 1436;
Best Local Similarity 41.7%; Pred. No. 7.9e-213;
Matches 613; Conservative 208; Mismatches 495; Indels 154; Gaps 23;

QY 28 CILLNSCFLLISFNGTDLERLVNGDPCSGRVEVVKFQGWGTVCDDGNNTASTVVK 87
Db 13 CVLLLG-----TWVGQALERLKDGVHRCGRVEVVKHQEWGTVDGYRWTLDKASVVC 67
QY 88 QLCGPFSFAM--FRFGQALVRGKTLWLDVSCYGNESALWEQCH--REWGSHNCYHGE 141
Db 68 QLCGGAIGPFGAYFGPL--GPIWLLYTSCEGTSTVSDCEHNSIKDYRNDGYNHGR 124

QY 142 DGVNVCYGEANLGLRLVDGNNSCGRVEVVKFQGRWGTICDDGNNTASTVVKRQLGCPSS 201
Db 125 DAGVVCSSG----FVRLAGDGPSCGRVEVHSEAWIPVSDGNFTLATQAIICAEIGCCGKA 180
QY 202 FISSGVNSPAVALRPIWLDLILCOGNELALMNCRRHGNHDCSHNEDVTLYTCYDSSDLE 261
Db 181 VSVLGHFELFRESSAQWAEFEFCEGEPELWVCPVPGGTCCHHSGSAQVVCYSAYSEVR 240
QY 262 LELVCGTNRCAGRVLEKIOGRWGTVCHHKNNAADVVKOLGCGCTALHFAGLPHLOSGS 321
Db 241 L-MTNGSSQCEGOVEMNISGOWRALCACSHWSLANANVICRQLGCGVAISTPGPHLYEBG 299
QY 322 DVMWLDGVSCSNESFLMDCRHSGTVNFDCLEHQNQVSVICSDGADLELRADGNNCSGR 381
Db 300 DQILARFHCSCAESFLWSCPVLTALGGPDCSHGNTASVICSGN----- 342
QY 382 VEVRIHEQWTTICDQNWKEQALVVKQLGCPFSVFGSRRAKPSNEARDIWNISCTGN 441
Db 343 -QIQVLPG-----CNDV-----SQTGSAAS----- 363
QY 442 ESALWDCITYDGKAKRTCFRRSDAGVICSDKADLRLVGAHSPCYGRLEVKYQGEWGTVC 501
Db 364 -----EDSAPYCSDSR--QLRLVDGGGPGCAGRVEILDQGSWGTIC 401
QY 502 HDRWSTRNAAVVKOLGCGKPMHFGMTYFKEASGPIWLDLDDVSCICGNESNIHDCRSHGW 561
Db 402 DGDWDLDDARVVCROLGCGEALNATGSAHFAGAGSGPIWLDNLNCTGKESHVWRCSRSGW 461
QY 562 KENCVHREDIVITCGDAGTWLRLVGGSNCRSLRLEVFGQGRWGTVCDDGNNSKAAAVVC 621
Db 462 QHNCRHKQDAGVICS--EFLALRWVSEDOOCAGWLEVFYNGTWSVCRNPMEDITVSTIC 519
QY 622 SOLDCPSSIIIGMLGNAST----GYGKIWLDVSDGDESDLWSCNRSWGNNDCHSHED 677
Db 520 RQLGCGDS----GTLNSSVALREGFRPQWVDRIQCRKTDTSILWQCPSPMNTNSCPKEE 575
QY 678 VGVICSDASDMLRLVGGSSRCAGKVEVNVQGAVGITLCANGMNMIAEVVCRQLGCGSAI 737
Db 576 ATIWACDS--RQIRLVDDGGRCGRVEILDQGSWGTICDDRWDLDARVVCROLGCGEAL 633
QY 738 RVSREPHTERTLHILMSNGCTGGEASLWDCIRWEKQOTACHLNWEASLCSAHRQPL 797
Db 634 DATVSSFFGTGSGPTWLDDEVNCRGEESQVWRCPFWGVRQHNCNHQEDAGVICSGF--VRL 691
QY 798 VGADMEPCSRGVEVVKHADTWRSVCDSDFLHAANVLRCLELNCGDALISLSVDHFHKGNGLT 857
Db 692 AGDGPSCGRVEVHSEAWTPVSDGNFTLPTAQVICAELCCGKAVSVLGHMPFRESGQV 751
QY 858 WAEKFCQEGSETHLALCPIVQHPEDTCIHSREVGVVCSRYTIDVRLV--NGKSCQDGOVEIN 916
Db 752 WAEFPCDGGEPPELWSCPRVPCPGGTCILHSGAAQVVCVYVTEVQLMKNGTSCQEGOVEMK 811
QY 917 VLGHWSLCDTHWDPEDARVLCRQLSCGTALSTTGKYGKIGERSVRYWGHFRHCLGNESLL 976
Db 812 ISGRWALCASHWSLANANVVCROLGCGVAISTPGRPHLVEGGDQISTAOFHCSGAESFL 871
QY 977 DNMCTVLGAPCIGHNTVSVICTGSLTOPLPFLCLANVSDPYLSAVPEGSALICLEDKRL 1036
Db 872 WSCPVTALGGPDCSHGNTASVICSGNHTQVLPQCNDFLSQPAGSAASESSPYCSDSRQL 931
QY 1037 RLVDGSRCAGRVEIYHDFGWTICDDGWDLSDAHVVVKOLGCGVAFAFATVSAHFEGSG 1096
Db 932 RLVDGSGPGCGGRVEILDQGSWGTICDDWDLDDARVVCROLGCGEALNATGSAHFAGSG 991
QY 1097 PTWLDLNTCTGTHLWQCPSCRGWQHCHRHKEDAGVICSEFTALRLYSETETESACARL 1156
Db 992 PTWLDLNTCTGTHLWQCPSCRGWQHCHRHKEDAGVICSEFTALRLYSETETESACARL 1049
QY 1157 EYFNGTWSVGRNNTITTAIGIVCRQLCGGNGVVSAPLSKTCGPMWDDIOCPKTH 1216
Db 1050 EYFNGTWSVGRNNTITTAIGIVCRQLCGGNGVVSAPLSKTCGPMWDDIOCPKTH 1109
QY 1217 ISIWQCLSPAPWERRISSPAEETWITCE-----DRIRVRGDDTECSGRVE 1260

Db 1110 TSLWQCPGPMKYSKSCPEKEAYTSCGRRPKSCPTAAACTDRKRLRGDSECSGRVE 1169
QY 1361 IWHAGSWCTVCDSDWDLAEAVWCQQLGCGSALAALRDASFGCGTGTIWLDDMRCKGNES 1320
Db 1170 VWHNGSGTVCDDSWSLAEAVWCQQLGCGSALAALRDASFGCGTGTIWLDDMRCKGNES 1229
QY 1321 FLWDHAKPMQSQCGHKEDAGVRCSG-----QSLKSLNASSGHLALIL 1364
Db 1230 SLWDCVAEPWQSQCKDEADAGVRCSGVRTTPTTTAGTRTTSNLSLPGIFSLPGVLCJL 1289
QY 1365 SSIFGLLLLVLFILFVTCRWOKHPLRVSTRRRGSLENLFHEMETCLKREDPHGTR 1424
Db 1290 GSLFLVLVLVTLQRLW-RAER-----RALSSYEDALAEAVYEELDYLLTOREGLSP 1342
QY 1425 TS-DTTPNHGCEADSDTSLGLVLPASEATK 1453
Db 1343 DQMTDVPDENYDDAEVVPVPGTSPSQGNE 1372

RESULT 12

US-09-759-130B-448
; Sequence 448, Application US/09759130B
; Publication No. US2003002279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; FILE OF INVENTION: US
; FILE REFERENCE: MPI00-5350MIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 448
; LENGTH: 1436
; TYPE: PRT
; ORGANISM: Bovine
US-09-759-130B-448

Query Match 37.9%; Score 3082; DB 9; Length 1436;
Best Local Similarity 41.7%; Pred. No. 7.9e-213;

Matches 613; Conservative 208; Mismatches 495; Indels 154; Gaps 23;
QY 28 CILLNSCLFSSNGTDLRLVNGDPCGCTVBFKFGQMGTVCDGWNNTASTVYCK 87
Db 13 CVLLIG-----TWGGQALELRKDGVRCEGRVEVKHQEGWGTVDGYRWTLKDVYCR 67
QY 88 QLGCPFPSPAM---PRFQAVTRHGKIWLDDVSCYGNESALMECOH---REMGSHNCHYE 141
Db 68 QLGCAAGFPFGGAYFGLPGL---GPIWLLYTSCEGTSTVSDCHSNIKDKVRNDGYNHGR 124
QY 142 DVGVNCEANLGLRLVDGNNCSGRVEVKFQERWGTTCDDGNNLNTAAYVCRQLGCPSS 201
Db 125 DAGVCSG-----FVRLAGDGPGRVEVHSGEAWIPVSDGNFTLATAIICAEELGCKA 180
QY 202 FISSGVNPAVLRPIWLDDILCOGNELALWNCRRHGNHDCSHNEDVTITCYDSSDLE 261
Db 181 VSLGLHELFPRESSAQVWAEERCEGEPELWCPVPCPGTCHHSGSAQVVCSEVR 240
QY 262 LRLVGGTNRNCRGRVELKTQGRWGTVCVHHKNNAAADVVCQKQCGGTALHFAGLPHLSGS 321
Db 241 L-MTNGSSQCEGQVEMNISQWRALCASHWSLANANVCRQLGCGVAISTPGGPHLVEEG 299
QY 322 DVWLVDGVCSCGNSFLLWDCRHSGTVNFDCLHONDVSVICSDGADLELRDLGDSNNSGR 381
Db 300 DQILTARFHCSGAESFLWSCPVTALGGPDCSHGNTASVCSGN----- 342
QY 382 VEVRIHEQWMTICDNWKNQALVVCQKQCPESVFGSRRRAKPSNEARDIWNISICTGN 441
Db 343 -QIQVLPQ-----CNDV-----SQPTGSAAS----- 363
QY 442 ESALWDCTYDGAKRTCFRRSDAGVICSADKADLRLVLAHSPCYGRLEVYQGEWGTVC 501
Db 364 -----EDSAPYCSDSR--QLRLVDGPGPCAGRVEILLDQGSWGTC 401
QY 502 EDRWSTRNAVVCQKCGKPMHVFGMTYFKEASCPILWDDVSCIGNESNWDCEHSGW 561
Db 402 DQGDWLDLDDARVVCQKCGEALNATGSAHFGAGSGPIWLNLNCTGKESHVWRCPSRMG 461
QY 562 KHNVCVHRDVTVTCGSDATGLRLVGGNRCSGRLEVYFQGRWGTVCDDGNSKAAAVVC 621
Db 462 QNCRHKODAGVCS--EFLALRVVSEDOQACAGLEVYNGTWGVCNPNMEDITVSTIC 519
QY 622 SOLDCPSSIIGMIGLNAST-----GYGKIWLDDVSCDSDGLWSCNRSNMGNNCSHSD 677
Db 520 QQLGCGDS-----GTLNSSLVALREGFRPQWDRIOCRKTDLSLWQCPSPWNYNSCPKEE 575
QY 678 VGVICSASDMELRLVGGSSRCAGKVEVNVGAVGILCANGWGMNIAEVVCRQLECGSAI 737
Db 576 AYIWCADS--RQIRLVDGGGRCGRVEILLDQGSWGTCDDRDWLDLDDARVVCQKCGEAL 633
QY 738 RVSREPHTERTLHLSNSGCTGGEASLWDCIRWENQOTACHLNMEASLICSARHQRPL 797
Db 634 DATVSSPFGTSGPIWLDEVNCRGEESQVWRCPSPGWHRQHNCNHOEDAGVICSGF--VRL 691
QY 798 VGADMPGSGRVEVKHADTWRSVCDSDFSLSHAANVLCRELNCGDALSLSVGHFGKNGLT 857
Db 692 AGDGPSCGRVEVHSGEAWTPVSDGNFTLPTAQVTCABELGCGKAVSLGHMPPFRESQGV 751
QY 858 WAEKFCQEGSETHALALCPVOHPEDTCHSRVGVVVCSTYDRLV--NGKSCQDQVEIN 916
Db 752 WAEFRCDGGEPELWSCPVPVPCPGTCLHSGAAQVSVYTEVQLMKNGTSOCEQGVEMK 811
QY 917 VLGHWSLCTDHPEDARVLCRLSCGTALSTTGGKYIGERSVRVWGHFRHCLGNESLL 976
Db 812 ISGRWALCASHWSLANANVVCRLGCGVAISTPRGPHLVEGGDQISTAQFHCSGAESFL 871
QY 977 DNCQMTVLGAPPCIHGNTVSVICTGSLTQPLPCLANVSDPYLSAVPEGSALICLEDKRL 1036
Db 872 WSCPVTALGGPDCSHGNTASVICSNGHTQVLPQCNDLFLSQPAGSAASESSPYCSDSRQL 931
QY 1037 RLVDGDSRCAGRVEIYHDFWGTTCDDGWDLSDAHYVCQKLGCGVAFNATVSAHFEGSG 1096
Db 932 RLVDGGGPGCGRVEILLDQGSWGTCDDWLDLDDARVVCRLGCGEALNATGSAHFAGSG 991

Search completed: May 12, 2003, 06:47:08
Job time : 162 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 12, 2003, 10:35:25 ; Search time 254 Seconds
(without alignments)
7108.859 Million cell updates/sec

Title: US-09-759-130B-381

Perfect score: 8138

Sequence: 1 MMLPQNSWHIDGRCCCHQN.....CEDASDTSLGLVPASEATK 1453

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 783854 seqs, 621352466 residues

Total number of hits satisfying chosen parameters: 1567708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cn2.1/USPTO.spool/US09759130/runat_06052003.123555.23406/app_query.fasta.1.1607
-DB=Published_Applications_NA -QWFAST=FASTA -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09759130.@CGN.1.1.158.0runat_06052003.123555.23406
-NCPU=6 -ICPU=3 -NO_XLPY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published_Applications_NA:*
1: /cn2.6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cn2.6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cn2.6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cn2.6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cn2.6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cn2.6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cn2.6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cn2.6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cn2.6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cn2.6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cn2.6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cn2.6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cn2.6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cn2.6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8138	100.0	4359	US-10-042-431-10	Sequence 10, Appl
2	8138	100.0	4359	US-09-759-130B-380	Sequence 380, App
3	8138	100.0	4628	US-10-042-431-9	Sequence 9, Appl
4	8138	100.0	4628	US-09-759-130B-379	Sequence 379, App

5	3082	37.9	4308	9	US-10-042-431-79	Sequence 79, Appl
6	3082	37.9	4308	9	US-09-759-130B-449	Sequence 449, App
7	1142.5	14.0	4344	10	US-09-917-800A-1712	Sequence 1712, Ap
8	830	10.2	2562	9	US-09-147-947-3	Sequence 5, Appl
9	818	10.1	473	9	US-09-918-995-32438	Sequence 32438, A
10	722	8.9	2614	9	US-09-147-947-3	Sequence 5, Appl
11	691.5	8.5	1686	9	US-09-905-291A-147	Sequence 147, App
12	691.5	8.5	1686	9	US-09-902-853-147	Sequence 147, App
13	691.5	8.5	1686	9	US-09-907-824-147	Sequence 147, App
14	691.5	8.5	1686	9	US-09-907-841-147	Sequence 147, App
15	691.5	8.5	1686	9	US-09-904-011-147	Sequence 147, App
16	691.5	8.5	1686	9	US-09-906-742-147	Sequence 147, App
17	691.5	8.5	1686	9	US-09-906-838-147	Sequence 147, App
18	691.5	8.5	1686	9	US-09-907-613-147	Sequence 147, App
19	691.5	8.5	1686	9	US-09-907-942-147	Sequence 147, App
20	691.5	8.5	1686	9	US-10-227-884-113	Sequence 113, App
21	691.5	8.5	1686	9	US-09-904-820-147	Sequence 147, App
22	691.5	8.5	1686	9	US-09-904-859-147	Sequence 147, App
23	691.5	8.5	1686	9	US-09-909-204-147	Sequence 147, App
24	691.5	8.5	1686	9	US-10-230-163-113	Sequence 113, App
25	691.5	8.5	1686	9	US-09-904-786-147	Sequence 147, App
26	691.5	8.5	1686	9	US-09-906-646-147	Sequence 147, App
27	691.5	8.5	1686	9	US-09-906-700-147	Sequence 147, App
28	691.5	8.5	1686	9	US-09-902-903-147	Sequence 147, App
29	691.5	8.5	1686	9	US-09-903-749A-147	Sequence 147, App
30	691.5	8.5	1686	9	US-09-903-786-147	Sequence 147, App
31	691.5	8.5	1686	9	US-10-218-631-113	Sequence 113, App
32	691.5	8.5	1686	9	US-10-230-338-113	Sequence 113, App
33	691.5	8.5	1686	9	US-09-902-736-147	Sequence 147, App
34	691.5	8.5	1686	9	US-09-904-119-147	Sequence 147, App
35	691.5	8.5	1686	9	US-09-904-956-147	Sequence 147, App
36	691.5	8.5	1686	9	US-09-907-794-147	Sequence 147, App
37	691.5	8.5	1686	9	US-10-230-414-113	Sequence 113, App
38	691.5	8.5	1686	9	US-09-902-692-147	Sequence 147, App
39	691.5	8.5	1686	9	US-09-903-520-147	Sequence 147, App
40	691.5	8.5	1686	9	US-09-903-943-147	Sequence 147, App
41	691.5	8.5	1686	9	US-09-905-056-147	Sequence 147, App
42	691.5	8.5	1686	9	US-09-907-925-147	Sequence 147, App
43	691.5	8.5	1686	9	US-09-904-553-147	Sequence 147, App
44	691.5	8.5	1686	9	US-09-905-381-147	Sequence 147, App
45	691.5	8.5	1686	9	US-09-905-381-147	Sequence 147, App

ALIGNMENTS

RESULT 1
US-10-042-431-10
; Sequence 10, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; FILE REFERENCE: 10147-602
; CURRENT APPLICATION NUMBER: US/10/042,431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 10
; LENGTH: 4359
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-431-10

Alignment Scores:

Pred. No.: 0

Length: 4359

Score: 8138.00 Matches: 1453
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-759-130B-381 (1-1453) x US-10-042-431-10 (1-4359)

QY 1 MetMetLeuProGlnAsnSerTrpHisIleAspPheGlyArgCysCysHisGlnAsn 20
 DB 1 ATGATGCTGCCTCAAACTCGGCGCATATGATTTTGAAGATGCTGCTCATCAGAAC 60
 QY 21 LeuPheSerAlaValThrCysIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuSerCysPheLeuIleSer 40
 DB 61 CTTTCTCTGCTGCTAACTTGCATCTGCTGCTCCTCAATTCCTGCTTCTCATCAGCAGT 120
 QY 41 PheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr 60
 DB 121 TTTAATGGAAACAGATTGGAGTTGAGGCTGGTCAATGGAGAGCGGTCCCTGCTCTGGGACA 180
 QY 61 ValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThr 80
 DB 181 GTGGAGGTGAATTCACGGACAGTGGGGACTGTGTGTGATGATGGGTGGAACACTACT 240
 QY 81 AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArgPhe 100
 DB 241 GCCTCAACTGCTGCTGCAACAGCTTGGATGCTCCATTTCTTTCGCCATGTTTCGTTTT 300
 QY 101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsn 120
 DB 301 GGACAAAGCCGTGACTAGACATGGAATAATTTGGCTTGATGATGTTCTCTTATGGAAT 360
 QY 121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140
 DB 361 GAGTCAGCTCTCTGGGAATGTCAACACGGGAATGGGAAGCAATACTTTATCATGGA 420
 QY 141 GluAspValGlyValAsnCysTyrGlyGluAlaAsnLeuLeuGlyLeuArgLeuValAspGly 160
 DB 421 GAAGATGTTGGTGTGAACCTGTTATGGTGAAGCCAACTCTGGTGTGAGGCTAGTGGATGA 480
 QY 161 AsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCys 180
 DB 481 ACAACTCTGTTTCAAGGAGAGTGGAGTGAATTTCCAAGAAGGTGGGGACTATATGT 540
 QY 181 AspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer 200
 DB 541 GATGATGGGTGAACCTGAATCTGCTGCCGTGGTGTGCAGGCAACTAGGATGTCATCT 600
 QY 201 SerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAsp 220
 DB 601 TCTTTTATTTCTTCGGAGTGTGTAATAGCCCTGCTGATTCGCCCCCAATTTGGCTGGAT 660
 QY 221 AspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGly 240
 DB 661 GACATTTATGCCAGGGGAATCAGTTGGCACTCTGGAATTCAGACATCGTGGATGGGA 720
 QY 241 AsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeu 260
 DB 721 AATCATGACTGCAGTCACAAATGAGGATGTCACATTAACCTGTTATGATAGTAGTACTCT 780
 QY 261 GluLeuArgLeuValClyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGln 280
 DB 781 GAACAAAGCCTGTAGGTGGAACTAACCCCTGTATGGGGAGAGTAGAGCTGAAATCCAA 840
 QY 281 GlyArgTrpGlyThrValCysHisIleTrpAsnAsnAlaAlaAspValValCys 300
 DB 841 GGAAGTGGGGACCGTATGCCACCAATAGTGAACAATGCTGCACTGATGCTGATGCT 900
 QY 301 LysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGly 320
 DB 901 AAGCAGTTGGGATGTGGAAACCCCACTTCACTTCCGTGGCTTCCTCAATTTGCACTGAGG 960
 QY 321 SerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAsp 340

DB 961 TCTGATGTTGATGGCTTCATGGTCTCCTGCTCCGGTAAATGAATCTTTCTTTGGGAC 1020
 QY 341 CysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIle 360
 DB 1021 TGCAGACATTCGGGAACCGTCAATTTTTCACCTGCTTCATCAAAACGATGTGTCTGTGATC 1080
 QY 361 CysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGly 380
 DB 1081 TGCTCAGATGGAGCAGATTTGGAACCTGGACTAGCAGATGGAAGTAAACAATTTGTTTCAGGG 1140
 QY 381 ArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsn 400
 DB 1141 AGAGTAGAGGTGAGAATTCATGAACAGTGGTGGACAATATGTGACCAACATGGAAGAAAT 1200
 QY 401 GluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArg 420
 DB 1201 GAACAAGCCCTTGTGTTTGAAGCAGCTAGGATGCTCGTTCAGCTCTTTGGCAGTCTGT 1260
 QY 421 ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly 440
 DB 1261 CGTGTAAACCTAGTAATGAAGCTAGACACATTTGGATAAACACAGCATATCTTGCACCTGG 1320
 QY 441 AsnGluSerAlaLeuTrpAspCysThrTrpAspGlyLysAlaLysArgThrCysPheArg 460
 DB 1321 AATGAGTCAGCTCTCTGGGACTGCAATATGATGGAAGCAAGCAAGCAATGCTTCCGA 1380
 QY 461 ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGly 480
 DB 1381 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTGCGG 1440
 QY 481 AlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrVal 500
 DB 1441 GCTCATAGCCCTGTTATGGGAGATTGGAGTGAATACCAAGGAGAGTGGGGAGCTGTG 1500
 QY 501 CysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGly 520
 DB 1501 TGTCATGACAGATGGAGCAAGGAATGCGCTGTGTGTGTGTAACAATTTGGATGTGGA 1560
 QY 521 LysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeu 540
 DB 1561 AAGCCTATGATGCTGTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTG 1620
 QY 541 AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp 560
 DB 1621 GATGACGTTTCTTGGATTGGAAATGAGTCAAAATATCTGGGACTGTCAACACAGTGGATGG 1680
 QY 561 GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr 580
 DB 1681 GGAAAGCATAAATGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGTCAATGCAACA 1740
 QY 581 TrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPhe 600
 DB 1741 TGGGCGCTGAGGCTGGTGGCGGAGCAACCGCTGCTCGGGAAGACTTGGAGGTGACTTT 1800
 QY 601 GlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValVal 620
 DB 1801 CAAGGACCGTGGGGCACAGTGTGTGATGACCGCTGGAACAGTAAAGCTGCACGCTGGGTG 1860
 QY 621 CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr 640
 DB 1861 TGTAACCGCTGAGCTGCCATCTTCTATCATTTGCGATGGGTCTGGGAAACGCTTCTACA 1920
 QY 641 GlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrp 660
 DB 1921 GGATATGAAAAAATTTGGCTCGATGATGTTTCTGCTGATGGAGATGAGTCAATCTCTGG 1980
 QY 661 SerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyVal 680
 DB 1981 TCATCAGGAACAGTGGGTGGGGAATAATGACTCAGTCAAGTCAAGATGTTGAGTG 2040
 QY 681 IleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAla 700

Db 2041 ATCTGTTCTGATCGATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGCTGTGCT 2100
QY 701 GlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGly 720
Db 2101 GGAAGATTGAGGTGAATGTCAGGGTGGCCGTGGGAATCTGTGCTAATGGCTGGGA 2160
QY 721 MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer 740
Db 2161 ATGAACATTGCTGAAGTTGTTTTCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCC 2220
QY 741 ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr 760
Db 2221 AGAGGCTCATTTACAGAAAGAAATACACATCTTAATGTGCAATCTGGCTGCACT 2280
QY 761 GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis 780
Db 2281 GGAGGGGAAGCCCTCTCTCGGATTGTATACGATGGAGTGGAAACAGAGCTGGCTGCAT 2340
QY 781 LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAla 800
Db 2341 TTAATATGAAGCAAGTTTGATCTGCTCAGCCACAGCCAGCCAGGCTGGTTGGAGCT 2400
QY 801 AspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys 820
Db 2401 GATATGCCCTGCTCTGGAGCTGTTGAAGTGAACATGACAGACATGGCCCTCTGTCTGT 2460
QY 821 AspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp 840
Db 2461 GATTCTGATTTCTTCTTCATGCTGCCAATGTCTGTGCAGAGAAATTAATTTGGAGAT 2520
QY 841 AlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu 860
Db 2521 GCCATATCTTCTTGTGGAGATCACTTTGGAAAAGGAATGCTCTAACTTGGGCCGAA 2580
QY 861 LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro 880
Db 2581 AGTTCCAGTGTGAAGGAGTGAACATCACTTGCATTATGCCCATTTGTTCAACATCCG 2640
QY 881 GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTrpThrAspVal 900
Db 2641 GAAGACACTTGTATCCACACAGAGAAAGTTGGAGTTGCTGTTCCCGATATACAGATGTC 2700
QY 901 ArgLeuValAsnGlyLysSerGlnCysAspGlyClnValGluIleAsnValLeuGlyHis 920
Db 2701 CGACTTGTGAATGGCAATCCAGTGTGACGGCAAGTGGAGATCAACGTTGTTGGACAC 2760
QY 921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940
Db 2761 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAGATGCCGTGTTCTATGCAGACAG 2820
QY 941 LeuSerCysGlyThrAlaLeuSerThrThrGlyLysTrpIleGlyGluArgSerVal 960
Db 2821 CTCAGCTGTGGACTGCTCTCAACACAGAGGAGAAATATATTGGAGAAAGAGTGT 2880
QY 961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGln 980
Db 2881 CGTGTGTGGGACACAGGTTTCAATGTTAGGGAATCAGTCACTTCTGGATAACTGTCAA 2940
QY 981 MetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThr 1000
Db 2941 ATCAGAGTTCTTGGAGCACCCTCTGTCATCCATGGAAATACGTCTGTGATCTGCACA 3000
QY 1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSer 1020
Db 3001 GGAAGCTTGACCCAGCCACCTGTTTCCATGCTCGCAATGTATCTGACCCATATTGCTCT 3060
QY 1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAsp 1040
Db 3061 GCAGTTCCAGAGGGCAGTGTCTTGTGATCTGTAGAGGACAAACGGCTCCCGCTAGTGGAT 3120
QY 1041 GlyAspSerArgCysAlaGlyArgValGluIleThrHisAspGlyPheThrGlyThrIle 1060
Db 3121 GGGGACAGCCGCTGTGCGGGGAGATAGATCTATCACAGCGGCTTCTGGGACCAACATC 3180

QY 1061 CysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080
Db 3181 TGTGATGACGGCTGGGACCTGAGCGATGCCCGTGTGTCTCAAAAGCTGGGCTGTGA 3240
QY 1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeu 1100
Db 3241 GTGGCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTGAGGCCCATCTGGCTG 3300
QY 1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120
Db 3301 GATGACCTGAACATGCACAGGAAGAGTCCCACTTGTGGCAGTGCCTTCCCGGGCTGG 3360
QY 1121 GlyClnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAla 1140
Db 3361 GGCAGCACGACATGCAGGCACAAGGAGGAGCAGGGGTCTCTGCTCAGAAATTCACAGCC 3420
QY 1141 LeuArgLeuTrpSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyr 1160
Db 3421 TTGAGGCTCTACAGTGAACCTGAAACAGAGAGCTGTCTGGGAGATTGGAAGTCTTCTAT 3480
QY 1161 AsnGlyThrTrpGlySerValGlyArgAsnIleThrThrAlaIleAlaGlyIleVal 1180
Db 3481 AAGCGGACCTGGGGCAGCGTGGCAGGAGGAACATCACACAGCATAGCAGGATTTGTG 3540
QY 1181 CysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThr 1200
Db 3541 TGCAGGCAGCTGGGCTCTGGGAGAAATGGAGTTGTCAGCCCTCGCCCTTATCTAAGACA 3600
QY 1201 GlySerGlyPheMetTrpValAspIleGlnCysProLysThrHisIleSerIleTrp 1220
Db 3601 GGCTCTGTTTCTATGTTGGTGGATGACATTCAGTGTCTTAAACGCATATCTCCATATGG 3660
QY 1221 GlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIle 1240
Db 3661 CAGTGCCTGTCTGCCCATGGGACGGAAGAAATCTCCAGCCCGCAGGAGAGACCTGGATC 3720
QY 1241 ThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260
Db 3721 ACATGTGAAGATAGAATAAGAGTGGTGGAGGAGACACCGAGTGTCTGGGAGAGTGGAG 3780
QY 1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280
Db 3781 ATCTGGCAGCAGGCTCTCTGGGACAGTGTGTGATGCTCTGGGACCTGGCGGAGGCG 3840
QY 1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAlaSer 1300
Db 3841 GAAGTGTGTGTACAGAGCTGGGCTGTGGCTCTCTGCTGCTGGCTGGCGGAGGAGCTTCG 3900
QY 1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320
Db 3901 TTTGGCCAGGGAACCTGGAACCATCTGTTGGATGCATGCGGTGCAAGGAAATGAGTCA 3960
QY 1321 PheLeuTrpAspCysHisAlaLysProTrpGlyClnSerAspCysGlyHisLysGluAsp 1340
Db 3961 TTTCTATGGGACTCTACGCAACCCCTGGGACAGAGTGTGCTGTGGAGGAGGAGAGAT 4020
QY 1341 AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeu 1360
Db 4021 GCTGGCTGAGGTGCTCTGGACAGTCTGTAATGCTGAAATGCTCTCAGGTCAATTTA 4080
QY 1361 AlaIleLeuLeuSerSerIlePheGlyLeuLeuLeuValLeuPheIleLeuPheLeu 1380
Db 4081 GCATTTATTTATCCAGTATCTTTGGGCTCTCTCTCTGTTGTTTATTTATTTCTTC 4140
QY 1381 ThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArg 1400
Db 4141 ACGTGTGCGGAGTTCAAAAAACAAATCTGCCCCCTCAGAGTTTCAACCAAGAGGAGG 4200
QY 1401 GlySerLeuGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspPro 1420
Db 4201 GGTTCCTCGAGGAGAAATTTATTCATGAGATGGAGACCTGCCCTCAAGAGAGAGACCCA 4260

Db 1141 AGACTAGAGGTGAGAAATTCATGAACAGTGGTGGCAATATATGTACCAGAACTGGAAGAAAT 1200
Qy 401 GluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArg 420
Db 1201 GAACAAGCCCTTGGTGGTTGTAAGCAGCTAGGATGTCGGTTCAGCGTCTTTGGCAGTGGT 1260
Qy 421 ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly 440
Db 1261 CGTCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACACCATATCTTGCACCTGGG 1320
Qy 441 AsnGluSerAlaLeuTrpAspCysThrTyrrAspGlyLysAlaLysArgThrCysPheArg 460
Db 1321 AATGAGTCAGCTCTCTGGGAGTCACACATATGATGAAAGCAAGCAACATGCTTCCGA 1380
Qy 461 ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGly 480
Db 1381 AGATCAGATGCTGGAGTAATTTGTCGTAAAGCAGATCTGGACCTTAAGGCTTGTGGG 1440
Qy 481 AlaHisSerProCysTyrrGlyArgLeuGluValLysTyrrGlnGlyGluTrpGlyThrVal 500
Db 1441 GCTCATAGCCCTCTTATGGGAGATTGGAGTGAATACCAAGAGAGTGGGGACTGTG 1500
Qy 501 CysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGly 520
Db 1501 TGTCTATCAGAGATGGAGCACAAGAAATGCAGCTGCTGTGTAAACAATTTGGGATGTGA 1560
Qy 521 LysProMetHisValPheGlyMetThrTyrrPheLysGluAlaSerGlyProIleTrpLeu 540
Db 1561 AAGCCTATGATGTTGGTGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGGCTG 1620
Qy 541 AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp 560
Db 1621 GATCAGCTTCTTCATGTTGGAATGAGTCAATATCTGGGACTGTGAACACAGTGGATGG 1680
Qy 561 GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr 580
Db 1681 GGAAAGCATAATTTGTATACAGAGAGGATGTGATTTAAACCTGCTCAGGTGATGCAACA 1740
Qy 581 TrpGlyLeuArgLeuValGlyLysSerAsnArgCysSerGlyArgLeuGluValTyrrPhe 600
Db 1741 TGGGGCCTGAGGCTGGTGGGGCGGAGCAACCGCTGCTGGGAAGACTGGAGGTGTACTTT 1800
Qy 601 GlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValVal 620
Db 1801 CAAGGACGGTGGGCACAGTGTGTATGACGGCTGGACAGTAAAGCTGCAGCTGTGGTG 1860
Qy 621 CysSerGlnLeuAspCysProSerSerIleGlyMetGlyLeuGlyAsnAlaSerThr 640
Db 1861 TGTAGCCAGCTGGACTGCCATCTTATCATTTGGCATGGGTCTGGAAACGCTTCTACA 1920
Qy 641 GlyTyrrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrp 660
Db 1921 GGATATGGAAAAATTTGGCTCGATGATGTTTCCGTGTATGGAGATGAGTCAGATCTCTGG 1980
Qy 661 SerCysArgAsnSerGlyTyrrPglYAsnAsnAspCysSerHisSerGluAspValGlyVal 680
Db 1981 TCATGCAAGAACACTGGGTGGGGAATAATGACTGCAGTCACTGAGATGTTGGAGTG 2040
Qy 681 IleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAla 700
Db 2041 ATCTGTCTGTATGTCATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTCT 2100
Qy 701 GlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGly 720
Db 2101 GGAAAACTGTAGGTGAATGTCCAGGGTCCCGTGGGAATTTCTGTGTCTAATGGCTGGGGA 2160
Qy 721 MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer 740
Db 2161 ATGACATGCTGAAGTTGTTTGGAGGCAACTTGAATGGGGTCTGCAATCAGGGTCTCC 2220
Qy 741 ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr 760

Db 2221 AGAGAGCTCATTTTACAGAAAGAACATTTACACATCTTAATGTGCAATTTCTGGTGCAC 2280
Qy 761 GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis 780
Db 2281 CGAGGGAAGCCTCTCTCTGGGATTTATACGATGGGAGTGGAAACACAGCTCGCTGTCT 2340
Qy 781 LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAla 800
Db 2341 TTAATATGGAAGCAAGTTTGATCTGCTACGCCACAGCAGCCAGCGCTGTTGGAGCT 2400
Qy 801 AspMetProCysSerSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys 820
Db 2401 GATATGCCCTGCTCTGGACGTTTGAAGTGAACATGTCAGACACATGCGCTCTGTCTGT 2460
Qy 821 AspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp 840
Db 2461 CATCTCATTTCTCTCTTCATGCTGCCANTGTGTGTGCAGAGAAATTAATTTGGGAGAT 2520
Qy 841 AlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu 860
Db 2521 GCCATATCTCTTTCTGGGAGATCACTTTGGAAGGGAATGGTAACTTGGGCCCAA 2580
Qy 861 LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro 880
Db 2581 AAGTTCCAGTGTGAAGGAGTGAACCTCACTTGCATTTATGCCCATTTGTTCAACATCCG 2640
Qy 881 GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrrThrAspVal 900
Db 2641 GAACACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTTCCGATATACAGATGTC 2700
Qy 901 ArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHis 920
Db 2701 CGACTTGTGAATGTCAAATCCAGTGTGACGGCAAGTGGAGATCAACGCTGCTGGACAC 2760
Qy 921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940
Db 2761 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAAGATGCCGTGTTCTATGCAGACAG 2820
Qy 941 LeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrrIleGlyGluArgSerVal 960
Db 2821 CTCAGCTGTGGGACTGTCTCTCAACCCAGAGAGAAATATATTGGAAAGAAGTGT 2880
Qy 961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGln 980
Db 2881 CGTGTGGGGACACAGGTTTCAITGCTTAGGAATGAGTCACCTTCGATAACTGTCAA 2940
Qy 981 MetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThr 1000
Db 2941 ATGACAGTCTTGGAGCACCTCCCTGTATCCATGCGAAATACTGTCTCTGTGATCTGCACA 3000
Qy 1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrrLeuSer 1020
Db 3001 GGAAGCCTGACCCAGCCACTGTTTCCATGCGCTCGCAATGATCTGACCCATATTGTCT 3060
Qy 1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAsp 1040
Db 3061 GCAGTTCAGAGGCGAGTGTCTTGTATCTGTAGAGGACAAACGCTCCGCTAGTGGAT 3120
Qy 1041 GlyAspSerArgCysAlaGlyArgValGluIleTyrrHisAspGlyPheTrpGlyThrIle 1060
Db 3121 GGGGACAGCCCTGTGCGGGAGAGTAGAGATCTATCAGACGCGCTTCTGGGGCACCAC 3180
Qy 1061 CysAspAspGlyTyrrAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080
Db 3181 TGTGATGACGGCTGGGACCTGAGCGATGCCACGTTGGTGTGTCAAAAGCTGGGTGTGGA 3240
Qy 1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeu 1100
Db 3241 GTGCCCTCAATGTCACCGTCTCTGCTCACTTTGGGAGGGGTGAGGGCCCATCTGGGTG 3300
Qy 1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120
Db 3301 GATGACCTGAACTGCACAGGAACGAGTCCCACTTGTGGCAGTGCCTTCCCGGCGCTGG 3360

Db 628 TCTTTTATTTCTTCTGGAGTTGTTAAATAGCCCTGCTCTATTGGCCCTCATTTGGCTGGAT 687
Qy 221 AspileLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGly 240
Db 688 GACATTTATGCCAGGGAATGAGTTGGCACTCTGGAAATTCGACACATCGTGGATGGGA 747
Qy 241 AsnHisAspCysSerHisAsnGluAspValThrLeuThrCystYrAspSerSerAspLeu 260
Db 748 AATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTATGATAGTAGTATCT 807
Qy 261 GluLeuArgLeuValGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGln 280
Db 808 GAACAAAGGCTGTAGGTGGAACCTAAACCGCTGTATGGGAGAGTAGAGCTGAAAAATCAA 867
Qy 281 GlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAspValValCys 300
Db 868 GGAAGGTGGGGGACCGTATGCCACCAATAGTGAACATGCTGCAGCTGATGCTGTATGC 927
Qy 301 LysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGly 320
Db 928 AAGCAGTTGGGATGTCGACCGCACTTCACCTCGCTGGCTTGCCTCATTTGCCAGTCAGGG 987
Qy 321 SerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAsp 340
Db 988 TCTCATGTTGTATGGCTTGTATGGTGTCTCTGCTCGGTAAATGATCTTTCTTTTGGGAC 1047
Qy 341 CysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIle 360
Db 1048 TGCAGACATTCGGGAACCGTCAATTTTGACTGTCTTCATCAAAACGATGCTGTGTATC 1107
Qy 361 CysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGly 380
Db 1108 TGCTCAGATGGAGCAGATTTGGAACCTGGCACTACAGATGGAAGTGAACATTTGTCAGG 1167
Qy 381 ArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsn 400
Db 1168 AGACTAGAGTGTAGAATTCATGAACAGTGGTGGCAATATGTGACCAGAACTGGAAGAA 1227
Qy 401 GluGlnAlaLeuValCysLysGlnLeuGlyCysProPheSerValPheGlySerArg 420
Db 1228 GAACAAGCCCTTGTGGTTGTAAACAGCTAGGATGTCGTTACCGCTCTTTGGCAGTCGT 1287
Qy 421 ArgAlaLysProSerAsnGluAlaArgAspileTrpIleAsnSerIleSerCysThrGly 440
Db 1288 CGTCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACACATATCTTGCACATGG 1347
Qy 441 AsnGluSerAlaLeuTrpAspCysThrTrpAspGlyLysAlaLysArgThrCysPheArg 460
Db 1348 AATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAGAAAGCAAGCAATGCTTCCGA 1407
Qy 461 ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGly 480
Db 1408 AGATCAGATGCTGGAGTAATTTGTTCATGAAGCAGATCTGGACCTTAAGGCTTGTCCGG 1467
Qy 481 AlaHisSerProCysTrpGlyArgLeuGluValLysTrpGlnGlyGluTrpGlyThrVal 500
Db 1468 GCTCATAGCCCTCTGTATGGGAGATTGGAGTGAAATACCAAGAGAGTGGGGACATGTG 1527
Qy 501 CysHisAspArgTrpSerThrArgAsnAlaValValCysLysGlnLeuGlyCysGly 520
Db 1528 TGTCTACAGATGAGCAGCAAGGAATGCACTGTGTGTGTGAACAATTTGGGATGTGGA 1587
Qy 521 LysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeu 540
Db 1588 AAGCCTATGCATGTGTTGTATGACCTATTTTAAAGAGCATCTCAGGACCTATTTTGGCTG 1647
Qy 541 AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp 560
Db 1648 GATCAGCTTTCTTCATTTGGAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGG 1707
Qy 561 GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr 580
|||||

Db 1708 GGAAGCATAAATTTGTGTACACAGAGAGATGTGATTGTAACCTGCTCAGGTGATGCAACA 1767
Qy 581 TrpGlyLeuArgLeuValGlyYserAsnArgCysSerGlyArgLeuGluValYrPhe 600
Db 1768 TGGGCCCTGAGGCTGTGGGGCGGAGCAACCGCTGCTCGGGAACAGTGGAGGTGTACTTT 1827
Qy 601 GlnGlyArgTrpGlyThrValCysAspGlyTrpAsnSerLysAlaAlaValVal 620
Db 1828 CAAGCAGCGTGGGCGACAGTGTGTGATGACGCGTGAACACAGTAAAGCTGCAGCTGTGTG 1887
Qy 621 CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr 640
Db 1888 TGTAGCAGCTGGGACTGCCCATCTTCTATCATTTGCGATGGGTCTGGAAACGCTTCTACA 1947
Qy 641 GlyTrpGlyLysIleTrpLeuAspValSerCysAspGlyAspGluSerAspLeuTrp 660
Db 1948 GGATATGGAANAATTTGGCTCCATGATGTTTCTCTGATGAGAGATGAGTCAAGATCTCGG 2007
Qy 661 SerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyVal 680
Db 2008 TCATGCAGGAACACGTGGGTGGGAAATAATGACTGCAGTCACTGAGTGAAGATGTTGGAGTG 2067
Qy 681 IleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyYserSerArgCysAla 700
Db 2068 ATCTGTTCTGTATGCTCGATATGAGCTGAGGCTTGTGGGTGGAAGCAGCAGTGTGCT 2127
Qy 701 GlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGly 720
Db 2128 GGAAGTGTGAGTGAATGTCACAGGTGCGGTGGGAATTCGTGTCTTAATGGCTGGGA 2187
Qy 721 MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer 740
Db 2188 ATGAACATTTGCTGAAGTGTGTCAGGCAACTTGAATGCGGTCTGCAATCAGGCTCTCC 2247
Qy 741 ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr 760
Db 2248 AGAGAGCTCTATTCACAGAAGAACATACACATCTTAATGTCGAATTCGGCTGCGACT 2307
Qy 761 GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis 780
Db 2308 GGAGGGAAGCCTCTCTCTGGGATTGTATAGATGGGAGTGGAAACACACACTCGCTGTCTAT 2367
Qy 781 LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAla 800
Db 2368 TTAATATGGAAGCAAGTTTGTATCTGCTCAGCCACAGCCAGCCAGGCTGTTGGAGCT 2427
Qy 801 AspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys 820
Db 2428 GATATGCCCTCTCTGGACGTGTTGAAGTGAACATGACAGACATGCGCTCTCTCTGT 2487
Qy 821 AspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp 840
Db 2488 GATTCTGATTTCTCTCTTCATGCTGCCAATGTGCTGTCCAGAGAATTAATTTGGAGAT 2547
Qy 841 AlaIleSerLeuSerValIleYsPheHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu 860
Db 2548 GCCATATCTCTTCTCTGGAGATCACTTTGGAAGGGAATGCTTAACCTTGGGCCGAA 2607
Qy 861 LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro 880
Db 2608 AAGTTCAGTGTGAAGGGAGTGAACCTCACCTTGCAATATGCCCATTTGTTCAACATCCG 2667
Qy 881 GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTrpThrAspVal 900
Db 2668 GAAGACACTTGTATCCACAGCAGAGAACTTGGAGTTGTCTGTTCCGATATACAGATGTC 2727
Qy 901 ArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHis 920
Db 2728 CGACTTGTGAATGGCAATCCAGTGTACGCGCAAGTGGAGATCAACGCTGCTTGGACAC 2787
Qy 921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940
Db 2788 TGGGCTCACTGTGTGACACCCACTGGGACCCAGAAAGATGCCCGTGTCTTATGACAGACAG 2847

Qy	1301	PheGlyClnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer	1320
Db	3928	TTTGGCCAGGGAACCTGGAAACCACTCTGGTTGGATGACATGCGGTGCAAAAGGAAATGAGTCA	3987
Qy	1321	PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp	1340
Db	3988	TTTCTATGGAGACTGTCCGCCAAACCTGGGACAGAGTGACTGTGGACACACAGAGAGAT	4047
Qy	1341	AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeu	1360
Db	4048	GCTGGCGTAGGTCCTGGACAGTCGCTGAATCACTGAATGCCTCTCAGGTCAATTAT	4107
Qy	1361	AlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPheIleLeuPheLeu	1380
Db	4108	GCACATTATTTATCCAGTATCTTTGGGTCCTCTCCGGTGTCTGTTATTCTATTCTC	4167
Qy	1381	ThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArg	1400
Db	4168	ACGTGGTGCCGAGTTCAGAAACAAAAACATCTGCCCTCAGAGTTTCAACCAAGAGGAGG	4227
Qy	1401	GlySerLeuGluGlnAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspPro	1420
Db	4228	GGTTCTCTCAGAGAGAAATTTATTCATGAGATGGAGACCTGCCTCAAGAGAGAGAGCCCA	4287
Qy	1421	HisGlyThrArgThrSerAspAspThrProAsnHisGlyCysGluAspAlaSerAspThr	1440
Db	4288	CATGGGACAGACCTCAGATGCACCCCCACCACCATGGTTGTGAAGATGCTAGGCACACA	4347
Qy	1441	SerLeuLeuGlyValLeuProAlaSerGluAlaThrLys	1453
Db	4348	TCGCTGTGGAGTCTCTCTCGCTCTGAAGCCACAAA	4386

RESULT 4

US-09/759-130B-379
Sequence 379, Application US/09759130B
Publication No. US2003002279A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: McCarthy, Sean A
APPLICANT: Fraser, Christopher C
APPLICANT: Sharp, John D
APPLICANT: Barnes, Thomas S
APPLICANT: Kirst, Susan J
APPLICANT: Mackay, Charles R
APPLICANT: Myers, Paul S
APPLICANT: Leiby, Kevin R
APPLICANT: Wrighton, Nicolas
APPLICANT: Goodearl, Andrew
APPLICANT: Holtzman, Douglas A
TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, AND OTHER
TITLE OF INVENTION: USES.
FILE REFERENCE: MP100-5350NMIM
CURRENT APPLICATION NUMBER: US/09/759,130B
CURRENT FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: US 09/479,249
PRIOR FILING DATE: 2000-01-07
PRIOR APPLICATION NUMBER: US 09/559,497
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 09/578,063
PRIOR FILING DATE: 2000-05-24
PRIOR APPLICATION NUMBER: US 09/333,159
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: US 09/596,194
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 09/342,364
PRIOR FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: US 09/608,452
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/393,996
PRIOR FILING DATE: 1999-09-10
PRIOR APPLICATION NUMBER: US 09/602,871
PRIOR FILING DATE: 2000-06-23

;; PRIOR APPLICATION NUMBER: US 09/420,707
;; PRIOR FILING DATE: 1999-10-19
;; NUMBER OF SEQ ID NOS: 460
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 379
;; LENGTH: 4628
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-759-130b-379

Alignment Scores:

Pred. No.: 0 Length: 4628
Score: 8138.00 Matches: 1453
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-759-130b-381 (1-1453) x US-09-759-130b-379 (1-4628)

Qy	1	MetMetLeuProGlnAsnSerTrpHisIleAspPheGlyArgCysCysHisGlnAsn	20
Db	28	ATGATGCTGCTCAAACTCGGCATATTGATTTTGAAGATGCTGCTCATCACAAC	87
Qy	21	LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSer	40
Db	88	CTTTTCTGCTGTGTGTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	147
Qy	41	PheAsnGlyThrAspLeuLeuLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr	60
Db	148	TTTAATGGAACAGATTTGGAGTTGAGCTGCTCAATGGACGGCTCCCTGCTGGACA	207
Qy	61	ValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThr	80
Db	208	GTGGAGTGAATTCAGAGGACAGTGGGACCTGTGTGATGATGGTGGGAACACTACT	267
Qy	81	AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArgPhe	100
Db	268	GCCTCAACTGTGTGTGCAAAACAGCTGGATGCTCAATTTCTTCGCCCATGTTTCGTTT	327
Qy	101	GlyClnAlaValThrArgHisGlyLysIleTrpLeuAspValSerCysTyrGlyAsn	120
Db	328	GGCAAGCCGTGCTAGACATGGAATAATTTGGCTTGTATGATGTTTCTGTATGGAAAT	387
Qy	121	GluSerAlaLeuTrpGluCysGlnHisArgGlnTrpGlySerHisAsnCysTyrHisGly	140
Db	388	GAGTCACCTCTCTGGGAATGTCAACACCGGAATGGGAACCACTACTGTTATCATGGA	447
Qy	141	GluAspValGlyValAsnCysTyrGlyClnAlaAsnLeuGlyLeuValAspGly	160
Db	448	GAAGATGTTGGTGTGAACCTGTTATGTGTAAGCCCAATCTGGGTTTGGAGCTAGTGGATGGA	507
Qy	161	AsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCys	180
Db	508	AACAACTCCCTGTTCCAGGAGAGTGGAGTGAATAATTCAGAAAGAGTGGGAGCTATATGT	567
Qy	181	AspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer	200
Db	568	GATGATGGTGGAACTTGAATGCTGCTGCCGTGTGTCAGGCAACTAGGATGTCCTACT	627
Qy	201	SerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAsp	220
Db	628	TCCTTTTATTTCTTCTGGAGTTGTTAAATAGCCCTGCTGCTATTTGCCCCCAATTTGGCTGGAT	687
Qy	221	AspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGly	240
Db	688	GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGCAATTCGACACATCTGGATGGGGA	747
Qy	241	AsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeu	260
Db	748	AATCATGACTGCATCAACATGAGAGTGTACATTAATCTGTTATGATAGTAGTATCTT	807
Qy	261	GluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGln	280

Db	808	GAACCTAAGGCTTGTAGTGGAACTAAACCGCTGTATGGGAGAGTAGAGCTAAAAATCAA	867
Qy	281	GlyArgTrpGlyThrValCysHisHisIleTrpAsnAsnAlaAlaAspValValCys	300
Db	868	GGAAGTGGGGACCGTATGCCACCAAGTGAACCAATGCTGCAGCTGATGTCGTATGC	927
Qy	301	LysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGly	320
Db	928	AAGCAGTTGGGATGTGAACCGCACTTCACCTCGCTGGCTTGCCTCATTTCCAGTCAGG	987
Qy	321	SerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAsp	340
Db	988	TCTGATGTTGTATGGCTGTGATGCTCTCTCCGCTAATGATCTTTCTTTGGGAC	1047
Qy	341	CysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIle	360
Db	1048	TGCAGACATTCGGGAACCGTCAATTTTGACTGTCTTCATCAAAACGATGTCGTGATC	1107
Qy	361	CysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGly	380
Db	1108	TGCTCAGATGGAGCAGATTTGGAACCTGCACTAGCAGATGGAAGTAACAATTTGTCAGG	1167
Qy	381	ArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsn	400
Db	1168	AGAGTAGAGGTGAGATTCATGAACAGTGTGGACAATATGTGACCAGACACTGGAAGAAT	1227
Qy	401	GluClnAlaLeuValValCysLysGlnLeuCysProPheSerValPheGlySerArg	420
Db	1228	GAACAAGCCCTTGTGGTTGTGAACAGCTAGGATGTCGTTTCAGCGCTTTGGGAGTCGT	1287
Qy	421	ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly	440
Db	1288	CGTCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAACACGATATCTTGCACCTGG	1347
Qy	441	AsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArg	460
Db	1348	AATGAGTCAGCTCTCTGGGACTGCAATATGATGAAAGCAAGCAACATGCTTCCGA	1407
Qy	461	ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuValGly	480
Db	1408	AGATCAGATGCTGGAGTAATTTGTGTGAAGCAGATCTGGACCTTAAGCCTTGTCTGG	1467
Qy	481	AlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrVal	500
Db	1468	GCTCATAGCCCTCTTATGGGAGATTTGGAGCTGAAATACCAAGGAGAGTGGGGAGCTG	1527
Qy	501	CysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGly	520
Db	1528	TGTCATACAGATGAGCACAAAGAAATGCAGCTGTGTGTAAACAATTTGGGATGTGA	1587
Qy	521	LysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeu	540
Db	1588	AAGCCTATGATGTTGTTGTTGATGACCTATTTTAAAGAGCATCAGGACCTATTGGCTG	1647
Qy	541	AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp	560
Db	1648	GATCAGCTTTCTTCATTGGAAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATG	1707
Qy	561	GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr	580
Db	1708	GGAAAGCATTAATTTGTACACAGAGAGATGTTGTTGCTTAACCTGCTCAGGTGATGCAACA	1767
Qy	581	TrpGlyLeuArgLeuValGlyLysSerAsnArgCysSerGlyArgLeuGluValTyrPhe	600
Db	1768	TGGGGCTTGGGCTGGGGCGGCGAGCAACCGCTGCTGGGAAGACTGGAGGTGACTTTT	1827
Qy	601	GlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValVal	620
Db	1828	CAAGGACGGTGGGCGACAGTGTGTGATGACCGCTGGAACAGTAAGCTGACAGCTGGTG	1887
Qy	621	CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr	640

Db 1888 TGTAGCCAGCTGGAGCTGCCCATCTTCTATCATTTGGCATGGGTCTGGGAAACGGCTTCTACA 1947
Qy 641 GlyTyrGlyLysIleTrpLeuAspValSerCysAspGlyAspGluSerAspLeuTrp 660
Db 1948 GGATATGGAAAATTTGGCTCGATCATGTTCTCTGTGATGAGATGAGTCAGATCTCTGG 2007
Qy 661 SerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyVal 680
Db 2008 TCATGCAGGAACAGTGGGTGGGGAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTG 2067
Qy 681 IleCysSerAspAlaSerAspMetCileuArgLeuValGlyGlySerSerArgCysAla 700
Db 2068 ATCTCTTCTGATGATCGATGATGAGCTGAGGCTTGGGTGGAGCAGGAGTGTGCT 2127
Qy 701 GlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGly 720
Db 2128 GGAAAAGTTGAGTGAGTGTCCAGGTCCTGGGTGGGAATCTGTGTCTAATGCGTGGGA 2187
Qy 721 MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer 740
Db 2188 ATGAACATTTGCTGAAGTTGTTGGCAGGCACTTGAATGTGGTCTGCAATCAGGCTCTCC 2247
Qy 741 ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr 760
Db 2248 AGAGAGCCTCATTTTCACAGAAAGAACATTACACATCTTAATCTCGAATCTCGCTGCAC 2307
Qy 761 GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis 780
Db 2308 GGAGGGGAAGCCTCTCTCTGGATTGTATACGATGGAGTGGAAACACACTGCGTGTCTAT 2367
Qy 781 LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAla 800
Db 2368 TTAATAATGGAAGCAAGTTGATCTGCTCAGCCCAAGCCAGCCAGCTGTTGGAGCT 2427
Qy 801 AspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys 820
Db 2428 GATATGCCCTCTGCTGGAGCTGTTGAAGTGAACATGACAGACACATGCGCTCTGCTGT 2487
Qy 821 AspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp 840
Db 2488 GATTCGATTTCTCTCTCATGCTGCCAATGTGCTGTGCAGAGAATTAATTTGTGGAGAT 2547
Qy 841 AlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu 860
Db 2548 GCCATATCTCTTCTGTGGAGATCACCTTGGAAAAGGAATGGTCTAATCTGGGCCGAA 2607
Qy 861 LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro 880
Db 2608 AAGTTCCAGTGTGAAGGAGTGAACCTCACCTTGATATGCCCATTTGTTCAACATCG 2667
Qy 881 GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTrpThrAspVal 900
Db 2668 GAAGACACTGTATCCACAGCAGAGAAGTTGGAGTTGCTGTGTTCCGATATACAGATGTC 2727
Qy 901 ArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHis 920
Db 2728 CGACTGTGAATGGCAATCCCACTGTCACGGGCAAGTGGAGATCAACGTGCTGGACAC 2787
Qy 921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940
Db 2788 TGGGGCTCACTGTGTGACACCCACCTGGGACCCAGGAAGATGCCCGTGTCTATGCAGACAG 2847
Qy 941 LeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerVal 960
Db 2848 CTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATAATATGGAGAAAAGTGT 2907
Qy 961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGln 980
Db 2908 CGTGTGTGGGGACACAGTTCATTGCTTAGGGAATGAGTCACTTCTGGATAAAGTCAA 2967
Qy 981 MetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThr 1000
Db 2968 ATGACAGTTCTTGGAGCAGCTCCCTGTATCCATGGNAATACTGCTCTCTGTATCTGCACA 3027

Qy 1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrIleuSer 1020
Db 3028 GGAAGCCTGACCCAGCCAGCTGTTTCCATGCTCCAAATGATCTATGACCCCATATTTGTCT 3087
Qy 1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAsp 1040
Db 3088 GCAGTTCAGAGGCGAGTGTCTTGTATCTGCTTAGAGACAAACCGCTCCGCTAGTGGAT 3147
Qy 1041 GlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIle 1060
Db 3148 GGGACAGCCGCTGTCCGGGAGAGTAGAGATCTATCACGACGCTTCTGGGGCACCATC 3207
Qy 1061 CysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080
Db 3208 TGTGATACGCTGGGACCTGAGCGATGCCACCGTGTGTCTCAAAAGCTGGGCTGGGA 3267
Qy 1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeu 1100
Db 3268 GTGGCCTTCAATGCCAGGCTCTGCTCACTTTGGGAGGGGTGAGGGCCCATCTGGCTG 3327
Qy 1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120
Db 3328 GATGACCTGAACCTGCACAGGAACGAGTCCCACTTGTGGCAGTCCCTTCCCGGCTGG 3387
Qy 1121 GlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAla 1140
Db 3388 GGGCAGCAGCACTGCAGGCACAGGAGCAGCGGCTCATCTGCTCAGAAATTCACAGCC 3447
Qy 1141 LeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyr 1160
Db 3448 TTGAGGCTCTACAGTGAACCTGAAACAGAGAGCTGTGCTGGGAGATGGAAGTCTTCAT 3507
Qy 1161 AsnGlyThrTrpGlySerValGlyArgArgAsnIleThrAlaIleAlaGlyIleVal 1180
Db 3508 AACGGACCTGGGACGCTCGCAGGAGGAGAACATCACACAGCATACAGCAGCATTTGTG 3567
Qy 1181 CysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThr 1200
Db 3568 TGCAGGCACTGGGCTGTGGGAGAAATGGAGTTGTACGCTCGGCCCTTTATCTAAGACA 3627
Qy 1201 GlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrp 1220
Db 3628 GGCCTGTGTTTCATGTGGGTGATGACATTCAGTGTCTTAAACGCATATCTCCATATGG 3687
Qy 1221 GlnCysLeuSerAlaProTrpGluArgIleSerSerProAlaGluGluThrTrpIle 1240
Db 3688 CAGTGCCTGTCTGCCCATGGGAGCAAGAAATCTCCAGCCAGCAGAGACCTGGATC 3747
Qy 1241 ThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260
Db 3748 ACATGTGAAGATAGAATAAGAGTGGTGGAGAGACACCGAGTCTCTGGGAGAGTGGAG 3807
Qy 1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280
Db 3808 ATCTGGCAGCAGGCTCTCTGGGGACAGTGTGTGATGACTCTCTGGGACCTGGCCGAGGG 3867
Qy 1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300
Db 3868 GAACTGGTGTGTACACAGCTGGGCTGTGGCTCTGCTGCTGGCTGCCCTGAGGACGCTTCG 3927
Qy 1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320
Db 3928 TTTGGCCAGGGAACCTGGAACCATCTGTTGGATGATGCGGTGCAAGGAAATGAGTCA 3987
Qy 1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340
Db 3988 TTTCTATGGGACTGTACGCCAAACCTGGGGACAGAGTACTGTGGACACAAAGAGAT 4047
Qy 1341 AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeu 1360
Db 4048 GCTGGCGTGGAGTGTCTGGACAGTCGCTGGAATCACTGAATGCCTCTCCTCAGTCAATTTA 4107

262	Db	262	---	GGCCCCATTTGGCTTTTGTATATCTCATGTGAAGGACAGAGTCAACTGTC	312	...
125	Qy	125	TrpGluCysGlnHis	-----ArgGluTrpGlySerHisAsnCysTyrHisGlyGlu	141	
313	Db	313	AGTCACTGTGAGCATCTCTAATAATAAGACTATCGTATGATGGCTATAATCACTGGTCGG	372		
142	Qy	142	AspValClyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspClyAsn	161		
373	Db	373	GATGCTGGAGCTAGTCTGCTCAGGA	420	----	
162	Qy	162	AsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAsp	181		
421	Db	421	GGACCTGTCTCAGGGGAGTAGAAGTGCATCTTGAGAGAAGCTTGCATCCAGTGTCTGAT	480		
182	Qy	182	AspGlyTrpAsnLeuAsnThrAlaAlaValCysArgGlnLeuGlyCysProSer	201		
481	Db	481	GGAACTTTCACACTCCCACTGCCAGATCATCTGTGCAGAGTGGGTGTGGCAGGCT	540		
202	Qy	202	PheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAsp	221		
541	Db	541	GTGTCGTCTGGACATGAGCTCTTCAGAGAGTCCAGTCCCGCAGGCTGGGCTGAAGAG	600		
222	Qy	222	IleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsn	241		
601	Db	601	TTCAGGTGTGAGGGGAGGAGCGCTGAGCTCTGGGTCTGCCCCAGAGTCCCTGTGCCAGG	660		
242	Qy	242	HisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuGlu	261		
661	Db	661	GGCAGTGTCCACCAGTGGATCTGCTCAGGTGTGTTTTCAGCATCTCAGAAGTCCG	720		
262	Qy	262	IleArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGly	281		
721	Db	721	CTC---ATGACAAACGGCTCCTCTCAGTCTGAAGGCGAGTGCAGATGAACATTTCTGGA	777		
282	Qy	282	ArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAspValValCysLys	301		
778	Db	778	CAATGGAGAGCGTCTGTGCTCCCACTCGAGTCTGGCCAATGCCAATGTTATCTGTCGT	837		
302	Qy	302	GlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGlySer	321		
838	Db	838	CAGCTCGGCTGTGGAGTTGCCATCTCCACCCCGGAGGACACACTTGGTGAAGAAGGT	897		
322	Qy	322	AspValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCys	341		
898	Db	898	GATCAGATCTTAACAGCCCGATTTCACCTGCTCTGGGGCTGAGTCTCTCTGTGGAGTTGT	957		
342	Qy	342	ArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCys	361		
958	Db	958	CCTGTGACTGCCCTGGGTGGTCTGCTACTGTTCCCATGGCAACACACGCTCTGTATCTGC	1017		
362	Qy	362	SerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArg	381		
1018	Db	1018	TCAGGAAC	1026	----	
382	Qy	382	ValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGlu	401		
1027	Db	1027	--CAGATCCAGGTGCTTCCCCAG	1047		
402	Qy	402	GlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArg	421		
1048	Db	1048	-----TGCAACGACTCCGTG	1062		
422	Qy	422	AlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrClyAsn	441		
1063	Db	1063	TCACAACCTACAGGCTCGCGCTCA	1089		
442	Qy	442	GluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArgArg	461		
1089	Db	1089	-----	1089		
462	Qy	462	SerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValClyAla	481		

Db 1090 GAGGACAGCGCCCTACTGCTCAGACAGCAGG-----CAGCTCCGCGCTGGTGAGCGG 1143
Qy 482 HisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCys 501
Db 1144 GCGCGTCCCTGCGCGGAGAGAGTGGAGATCCTTGACAGAGGCTCTCGGGGACCATCTGT 1203
Qy 502 HisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLys 521
Db 1204 GATGACGCGTGGACCTGACGATGCCCGCTGGTGTGTCAGGCAGCTGGGTGGAGAA 1263
Qy 522 ProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAsp 541
Db 1264 GCCCTCAATGCCACGGGCTGCTCACTTCGCGGCGAGGATCAGGCGCCCATCTGTGTGAC 1323
Qy 542 AspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGly 561
Db 1324 AACCTGAACGCACAGAAAGAGTCCACAGTGTGGAGGTGCCCTTCCGGGGCTGGGG 1383
Qy 562 LysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrp 581
Db 1384 CAGCACAACTGCAGACACAAAGCAGCGGGGTCTACTCTGCTCA-----GAGTTCTCTG 1437
Qy 582 GlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGln 601
Db 1438 GCCCTCAGATGGTGTGAGTGTGAGTGTGCTGTGGTGTGCTGGAAGTTTCTTACAA 1497
Qy 602 GlyArgTrpGlyThrValCysAspGlyTrpAsnSerLysAlaAlaValValCys 621
Db 1498 GGGACCTGGGCGAGTCTGCGTAACCCATGGAAGACATCACTGTGCCAGATCTGC 1557
Qy 622 SerGlnLeuAspCysProSerSerIleleGlyMetGlyLeuGlyAsnAlaSerThr--- 640
Db 1558 AGACAGCTTGGTGTGGGACAGT-----GGAACCCCTCAACTCTCTTGTCT 1605
Qy 641 -----GlyTyrGlyLysIleTrpLeuAspValSerCysAspGlyAspGluSer 657
Db 1606 CTTAGAGAAAGTTTAGGCCACAGTGGTGGTAGAATTAACAATCATGCTTCCAAAGGAGAA 1665
Qy 658 AspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAsp 677
Db 1666 TCTCTCTGCGAGTCTTCTGACCTTGAATTAACAATCATGCTTCCAAAGGAGAA 1725
Qy 678 ValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyCysSer 697
Db 1726 GCTATATCTGGTGTGTCAGACAGC-----AGACAGATCCGCTGGTGGATGGAGTGGT 1779
Qy 698 ArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsn 717
Db 1780 CCGTCTCTGGGAGAGTGGAGATCCTTGACAGGCTCTCGGGGACCATCTGTGATGAC 1839
Qy 718 GlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGlyCysGlySerAlaIle 737
Db 1840 CGCTGGGACCTGGAGATGCCCGTGTGTGTGTCAGAGCAGTGGGTGGAGAGCCCTG 1899
Qy 738 ArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSer 757
Db 1900 GAGCGCACTGCTCTCTCTCTCTCTGGGAGCGGATCAGGCGCCCATCTGGCTGATGAAGTG 1959
Qy 758 GlyCysThrGlyGlyAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThr 777
Db 1960 AACTGCAGAGAGAGAGTCCCAAGTATGAGGTGCCCTTCTCGGGGATGGCGGCAACAC 2019
Qy 778 AlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeu 797
Db 2020 AACTGCAATCATCAAGAGATCAGGAGTCACTGCTCAGGATTT-----GTGCGTCTG 2073
Qy 798 ValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArg 817
Db 2074 GCTGGAGAGATGGACCTCTGTCAGGCGCAGTAGAAGTGCATCTCGAGAGCCCTGGACC 2133
Qy 818 SerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsn 837
Db 2134 CCAGTGTCTGTGGAACCTTCACACTCCCATCTGCCAGGTCTCTGTGCAGAGCTGGGA 2193

Qy 838 CysGlyAspAlaIleSerLeuSerValCysAspHisPheGlyLysGlyAsnGlyLeuThr 857
Db 2194 TGTGGCAAGGCTGTCTCTCTGGGACACATGCCATTCAGAGATCGGATGGCCAGGTC 2253
Qy 858 TrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleVal 877
Db 2254 TGGGCTCAAGAGTTTCAGGTGTGATGGGGGAGGCTGTGAGCTCTGCTGCCCCAGAGTG 2313
Qy 878 GlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyr 897
Db 2314 CCCTGTCCAGAGGACACATGTCTCCACAGTGGAGCTGTCTGAGTGTCTGTCTGCTGATC 2373
Qy 898 ThrAspValArgLeuVal---AsnGlyLysSerGlnCysAspGlyGlnValIleLeuAsn 916
Db 2374 ACAGAACTCCAGCTTATGAAAAACGGCACCTCTCAATGTGAGGGGACGAGTGGAGATGA 2433
Qy 917 ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal 936
Db 2434 ATCTGTGACGATGGAGAGCGCTCTGTGCCCTCCACATGGAGTCTGGCCCAATGCTT 2493
Qy 937 LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGly 956
Db 2494 GTCGTCTCAGCTCGGCTGTGGAGTCCCATCTCCACCCAGAGGACACACTTGGTG 2553
Qy 957 GluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu 976
Db 2554 GAAGAGGTGATCAGATCTCAACAGCCCAATTCACCTGCTCAGGGGCTGAGTCTCTCTG 2613
Qy 977 AspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSer 996
Db 2614 TGGAGTTTCTGTGACTGCTTGGTGGGCTGACTGTCCCATGGCAACACACCTCT 2673
Qy 997 ValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAsp 1016
Db 2674 GTGATCTCTCAGAAACACACAGGCTGTGCCCCAGTGCACAGACTTCTCTGCTCAA 2733
Qy 1017 ProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeu 1036
Db 2734 CCTGAGGCTCTCGGCGCTCAGAGAGAGTTCTCCCTACTGCTCAGACAGCAGGAGCTC 2793
Qy 1037 ArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPhe 1056
Db 2794 CCGCTGTGGAGGGGGCGGCTCCCTCGCGGGGAGAGTGGAGATCTTTCACAGGCTCC 2853
Qy 1057 TrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLys 1076
Db 2854 TGGGCACTCTGTGATGATGACTGGGACCTGGAGATGCCCGTGTGGTGTGACGGCAG 2913
Qy 1077 LeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluClySerGly 1096
Db 2914 CTGGCTGTGGAGAGCCCTCAATGCCACGGGTCTGCTCACTTCGGGCGAGATCAGGG 2973
Qy 1097 ProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysPro 1116
Db 2974 CCGCTGTGGTGGACCATCTGACTGCACAGGAAGAGTCCACGCTGTGGAGGTGCCCT 3033
Qy 1117 SerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSer 1136
Db 3034 TCCCGGGCTGGGGCGGCAGCTGACACACAAAGGAGGACCGCGGGTCTCTCTCTCA 3093
Qy 1137 GluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeu 1156
Db 3094 GAGTTCTTGGCGCTCAGGATGGTGAGC-----GAGGACACAGCAGTGTCTGGGTGGCTG 3147
Qy 1157 GluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrAlaIle 1176
Db 3148 GAGGTTTCTACACAGCGGACCTGGGGCAGTGTCTCCCGCAGGCCCTCCATGGAATATCACT 3207
Qy 1177 AlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaPro 1196
Db 3208 GTGTCGCTGATCTGCAGACAGCTTGGATGTGGGGACAGTGGAGAGTCTCAACACCTCTGT 3267

```
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, AND OTHER THERAPEUTIC USES.
; FILE REFERENCE: MPI00-535OMNIM
; CURRENT APPLICATION NUMBER: US 09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 449
; LENGTH: 4308
; TYPE: DNA
; ORGANISM: Bovine
US-09-759-130B-449

Alignment Scores:
Pred. No.:      2,74e-302          Length:    4308
Score:           3082.00            Matches:     613
Percent Similarity: 55.85%         Conservative:   208
Best Local Similarity: 41.70%       Mismatches:   495
Query Match:      37.87%           Indels:       154
DB:                9               Gaps:         23

US-09-759-130B-381 (1-1453) x US-09-759-130B-449 (1-4308)

QY 28 CysIleLeuLeuAsnSerCysPheLeuIseSerPheAsnGlyThrAspLeuGlu 47
    |||::||| |||||| |             ::              -ACCATGGTGGGTCTCAAGCTCTGGAG 81
Db 37 TGTCCTCTCCCTCGGC-----ACCATGGTGGGTCTCAAGCTCTGGAG 81

QY 48 LeuArgValAlaAsnGlyAspGlyProCysSerGlyThrValGluValLysPheGlnGly 67
    ||||| |||| :||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 82 CTGAGGTTGAAGGATGAGTGCCATCGCTGTGAGGGGAGACTGGAAGTGAAGCACCAAGA 141

QY 68 GluTrpGlyValCysAspGlyTrpAsnGlyThrAlaSerThrValValCysLys 87
    ::: ||| ||||| ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 142 GAATGGGCCACAGTGATGTTACAGTGACATTGAAGGATGATCTGTAGTGTGCAGA 201

QY 88 GlnLeuCysProPheSerPheAlaMet-----PheArgPheGlyGlnAlaVal 104
    ||||| ||||| :||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 202 CAGCTGGGGTGGAGCTGCCATTGGTTTTCTGGAGGGGCTTATTTTGGCCCAGGACTT 261

QY 105 ThrArgHisGlyLysIleTrpLeuAspValserCystyrGlyAsnGluSerAlaLeu 124
    ||| ||||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
Db 262 -----GGCCCCAATTGGCTTTGTATACTTAATCTGATGTTGAAGGCACAGAGTCAACTGTC 312

QY 125 TrpGlucysGlnHis-----ArgGluTrpClysSerHisasnCystrHisGlyGlu 141
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 313 AGTGACTGTGAGCATCTTAATAATAAGACTATCGTAACTGATGGCTATAATCATGTGTGCG 372

QY 142 AspValGlyValAsnCysTyrvirGlvLuAlaAsnLeuGlcylvLeuarqLeuvalaspGivAsn 161
```

Db 373 GATGCTGGAGTAGTCTGCTCAGGA-----TTTGTGCGGTCTGGCTGGAGGGAT 420
 Qy 162 AsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAsp 181
 Db 421 GGACCTGCTCAGGCGAGTAGAAGTCATCTGGAGAAGCTTGATCCAGTGTCTGAT 480
 Qy 182 AspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSerSer 201
 Db 481 GGGAACTTTCACACTTGCACATGCCAGATCATCTGTCAGAGATTGGTGTGGCAAGCT 540
 Qy 202 PheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAspAsp 221
 Db 541 GTGCTGCTCTGGGACATGAGCTTTCAGAGAGTCCAGTGGCCAGGTCTGGCTGAAGAG 600
 Qy 222 IleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsn 241
 Db 601 TTCAGTGTGAGGGGAGGAGCTGAGCTGCTGCTGCCAGAGTCCCTGTCTCAGG 660
 Qy 242 HisAspCysSerHisAsnGluAspValThrLeuThrCysTrpAspSerAspLeuGlu 261
 Db 661 GGCAGTGTACACAGAGTGTCTGCTCAGGTTGTTTTCAGCATATCTCAGAAGTCCGG 720
 Qy 262 LeuArgLeuValGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGly 281
 Db 721 CTC----ATGACAAACGGCTCTCTCAGTGTGAAGGCGAGTGGAGATGAACATTTCTGGA 777
 Qy 282 ArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAspValValCysLys 301
 Db 778 CAATGGAGAGGCTGTGTGCTCTCCACATGGAGTCTGGCCAAATGTTATCTGTCTG 837
 Qy 302 GlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuLysSerGlySer 321
 Db 838 CAGCTCGGCTGTGGATTGGCATTCACCCCGGAGGACACACTTGGTGGAGAAGGT 897
 Qy 322 AspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCys 341
 Db 898 GATCAGATCCTAACAGCCGATTTTCACTGCTCTGGGGTGTGATCTCTCTGGAGTTGT 957
 Qy 342 ArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCys 361
 Db 958 COTGTGACTGCCCTGGGTGGTCTGACTGTTCCTCCATGGCAACACAGCCTCTGTGATCTGC 1017
 Qy 362 SerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArg 381
 Db 1018 TCAGAAAC----- 1026
 Qy 382 ValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGlu 401
 Db 1027 ---CAGATCCAGGTGCTTCCCGAG----- 1047
 Qy 402 GlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArg 421
 Db 1048 -----TCAACGACTCCGTG----- 1062
 Qy 422 AlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsn 441
 Db 1063 TCTCAACTACAGGCTCTGCGGCCTCA----- 1089
 Qy 442 GluSerAlaLeuTrpAspCysThrTrpAspGlyLysAlaLysArgThrCysPheArgArg 461
 Db 1089 ----- 1089
 Qy 462 SerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAla 481
 Db 1090 GAGGACAGCGCCCTACTGTCTCAGACAGCAGG-----CAGCTCCGCTGTGTGACCGG 1143
 Qy 482 HisSerProCysTrpGlyArgLeuValLysTyrgingGlyGluTrpGlyThrValCys 501
 Db 1144 GCGGCTCCTCGCGGAGAGTGGAGATCTTGCACAGGCTCTCTGGGCGACCATCTGT 1203
 Qy 502 HisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLys 521
 Db 521 ----- 521

Db 1204 GATGACGCTGGGACCTGGACGATCCCGCGTGTGTGTGTCAGGACGCTGGGTGTGGAGAA 1263
 Qy 522 ProMetHisValPheGlyMetThrTrpPheLysGluAlaSerGlyProIleTrpLeuAsp 541
 Db 1264 GCCCTCAATGCCACGGGTCTGCTACATTCGGGGCAGGATCAGGCCCATCTGGTTGGAC 1323
 Qy 542 AspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGly 561
 Db 1324 AACTTGAACCTGCACAGGAAGAGTCCACAGTGTGAGGTGCCCTTCCCGGGCTGGGG 1383
 Qy 562 LysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrp 581
 Db 1384 CAGCACACTGCAGACACAGCAGCGGGGGTCACTGCTCA-----GAGTCTCTG 1437
 Qy 582 GlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTrpPheGln 601
 Db 1438 GCCCTCAGAGTGTGAGTGAGCAGCAGCAGTGTGCTGGTGTGCTGGAAGTTTCTACAAT 1497
 Qy 602 GlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValValCys 621
 Db 1498 GGGACCTGGGCGAGTGTGCTGCTAACCCCATGGAAGACATCAGTGTGTCCAGCATCTGC 1557
 Qy 622 SerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr--- 640
 Db 1558 AGACAGCTTGGCTGTGGGACAGT-----GGACCTCACTTCTTCTTGTCT 1605
 Qy 641 -----GlyTrpGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSer 657
 Db 1606 CTTAGAGAAGTTTTAGGCCACAGTGGTGGATAGATCCAGTGTGGAAAACTCACAC 1665
 Qy 658 AspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAsp 677
 Db 1666 TCTCTCTGGCAGTGTCTTACCCCTTGGAAATTACAACTCATGCTCTCCAAAGGAGGAA 1725
 Qy 678 ValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSer 697
 Db 1726 GCCTATATCTGCTGTGCAGACAGC-----ACACAGATCCGCTGTGGATGGAGTGT 1779
 Qy 698 ArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsn 717
 Db 1780 CGCTGCTCTGGAGAGTGGAGATCCTTACCAGGGCTCTCTGGGGACCATCTGTGATGAC 1839
 Qy 718 GlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGlyCysGlySerAlaIle 737
 Db 1840 CGCTGGGACCTGGACCATGCCGTGTGTGTGCAAGCAGCTGGGTGTGGAGAAGCCCTG 1899
 Qy 738 ArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSer 757
 Db 1900 GACGCCACTGTCTCTCTCTTCTGGGACGGATCAGGGCCCATCTGCTGGATGAAGTG 1959
 Qy 758 GlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThr 777
 Db 1960 AACTGCAGAGAGAGGAGTCCCAAGTATGGAGTGCCTTCTGGGGATGGCGGCAACAC 2019
 Qy 778 AlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeu 797
 Db 2020 AACTGCATCATCAAGAAGATCGAGGAGTCACTCTCTCAGGATTT-----GTGCGTCTG 2073
 Qy 798 ValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArg 817
 Db 2074 GCTGGAGAGATGAGCCCTGCTCAGGGCGAGTAGAAGTGCATTCGGAGAAGCCCTGGACC 2133
 Qy 818 SerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsn 837
 Db 2134 CCAGTGTCTGTGAAGAACTTCACACTCCCCACTGCCAGGTCATCTGTGAGAGCTGGGA 2193
 Qy 838 CysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThr 857
 Db 2194 TGTGGCAAGGCTGTGTCTGCTCTGGGACACATGCCATTCAGAGATCCGATGGCCAGGTC 2253
 Qy 858 TrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleVal 877
 Db 2254 TGGGCTGAAGAGTTACAGTGTGTATGGGGGGGAGCGCTGAGCTCTGTGCTCTGCCAGAGTG 2313

PRIOR APPLICATION NUMBER: US 60/290,645
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: US 60/292,336
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/295,798
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/297,457
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,884
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,459
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 1740
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1712
LENGTH: 4344
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_022849
US-09-917-800A-1712

Alignment Scores:
Pred. No.: 178e-105 Length: 4344
Score: 1142.50 Matches: 331
Percent Similarity: 41.73% Conservative: 136
Best Local Similarity: 29.58% Mismatches: 343
Query Match: 14.04% Indels: 312
DB: 10 Gaps: 39

US-09-759-130B-381 (1-1453) x US-09-917-800A-1712 (1-4344)

QY 258 SerAspLeuGluLeuArgLeuValGlyThrAsnArgCysMetGlyArgValGluLeu 277
Db 10 TCTGTTTGGCTGCTAGGCTGGTGAATGAGGAGACAGGTGCGGGTCTGGGAGATC 69
QY 278 LysIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAsp 297
Db 70 CTTTACAGGGTCTCTGGGCGACCATGTGTGATGACAGCTGGGACATCAATGATGCCAAC 129
QY 298 ValValCysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeu 317
Db 130 GTGGTGTGAGGAGCTGGGCTGTGGCTGTGGCTTCTCTGCCCGAGGAGTCCCGAGTTT 189
QY 318 GlnSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPhe 337
Db 190 GGACAGGGTCTGGTCCCATTTCTGGATGACGTGGCTGTAGAGGACATGAGGCTAT 249
QY 338 LeuTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspVal 357
Db 250 CTGTGGAGCTGTCTCCACCGAGGCTGGCTCTCTCAATAACTGTGGACATCAGGAGGATGCT 309
QY 358 SerValIleCysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsn 377
Db 310 GGAGTGTATGCTTCAGAT-----TCTCAACACAGCATGCC 345
QY 378 CysSerGlyArgValGluValArgIleHisGluGlnTrpThrIleCysAspGlnAsn 397
Db 346 ACACCCGGT-----TGTTGG-----AACCCCGGG 369
QY 398 TrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPhe 417
Db 370 GGCACAAATAACGATGTGATCTACGACACTCAAGAACCC----- 408
QY 418 GlySerArgArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSer 437
Db 409 ---ACAGAACTTCTCAACACAGCATCTCCACACCTGATGTTGG----- 450
QY 438 CysThrGlyAsnGluSerAlaLeuTrpAspCysThrTrpAspGlyLysAlaLysArgThr 457
Db 451 -----AACCATGGGGGCACAAATAATGATGTGATCTATGAC-----ACTCAAGAA 495
QY 458 CysPheArgArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArg 477

Db 496 ACCACAGAAAGAACAGATCTGGT-----TTGGCTGTGAGG 531
QY 478 LeuValGlyAlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyLeuTrp 497
Db 532 CTGGTGAATGGAGAGACAGGTGTCGGGGTGTGTGGAGATCCTTTACAGGGTTCCTGG 591
QY 498 GlyThrValCysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeu 517
Db 592 GGCACCGTGTGTGACGACAGCTGGGACATCAATGATGCCAAGCTGTGTGACGAGCTG 651
QY 518 GlyCysGlyLysProMetHisValPheGlyMetThrTrpPheLysGluAlaSerGlyPro 537
Db 652 GGTGTGGCTGGGCTTGTCTGCCCGAGGAGTGGCCAGTTTGGACAGGGCTCTGGGTCC 711
QY 538 IleTrpLeuAspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHis 557
Db 712 ATGTCTTGGATGACGTGGCTGTAGAGGACATGAGGCTATCTCTGGAGCTGCTCCAC 771
QY 558 SerGlyTrpGlyLysHisAsnCysValHisArgGluAspValIleValThrCysSer--- 576
Db 772 CGAGGCTGGCTCTCTCAATACTGTGACATCAGGAGGATGCTGGAGTCTATCTGTTCATAT 831
QY 576 ----- 576
Db 832 TCTCAACACAGCAGTCCACACCCGATCTCTCAACACAGCAGTCCACACCCGGTGGTGG 891
QY 577 -----GlyAspAlaThrTrpGly----- 582
Db 892 AACCCCGGGGGCACAAATAACGATGTGTCTATGGACCGGACACAGACACGACGCAACA 951
QY 583 -----LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGlu 597
Db 952 GATTCGTGGTTCGGCTGTAGGCTGGTGAATGGAGGACAGGTGTCAGGCTCGTGGAG 1011
QY 598 ValTyPheGlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAla 617
Db 1012 ATCTTTTACAGGGTTCCTGGGGTACCGTGTGTGACGACAGCTGGGACACCAAGGATGCC 1071
QY 618 AlaValValCysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsn 637
Db 1072 AACGTGGTGTGACGAGCAGCTGTGTGGCTGGCTCTCTGCCCGAGGAGTGGCCAC 1131
QY 638 AlaSerThrGlyTyGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSer 657
Db 1132 TTTGGACAAAGGCTCTGGATCCATTTCTGGATGACGTGGCTGTACAGGACATGAGGC 1191
QY 658 AspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAsp 677
Db 1192 TATCTGTGGAGCTGTCCCGGAGGCTGGCTCTCTCAATACTGTGGCCACCATGAGGAT 1251
QY 678 ValGlyValIleCysSerAspAlaSer-----AspMet----- 688
Db 1252 GCTGGAGTCTATGTTTCAGATGCCCAACCCAGAGCACAACTGGCCAGATATGTGGCT 1311
QY 689 -----Glu-LeuArgLeuValGlyGlySerSerArgCysAlaGlyLysValG 704
Db 1312 ACTACCACTCCAGAACTACAACAGATGTGGGACT-ACAAAATATCTCTCTCTCTCTCC 1370
QY 704 vAlaAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsnIleAl 724
Db 1371 TACAACACAAATTCGCCACCATTA-----GCCGATTGGTGGCAACAACCTCTCTCC 1418
QY 724 aGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer----- 740
Db 1419 GGAATACACC-----TGTGGAGGTTTACTGACCCCTACCCCTATGGCAGTT 1463
QY 741 -ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerClyCysTh 760
Db 1464 TTCCAGCCCATCATACCTCCGGAAGC-----TATCCTTAACAATGCCAGATGT--- 1509
QY 760 rGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysH 780

Db 1510 -----TTGTGG----- 1515
Qy 780 sLeuAsnMetGluAlaSerLeuLeuCysSerAlaHisArgGlnProArgLeuValGlyAl 800
Db 1516 -----AAATTTTCGTCTCCAGCATGAACCGTGTACAGAGTGCTTCAC 1559
Qy 800 aAspMetProCysSerArgValGluValLysHisAlaAspThrTrpArgSerValCy 820
Db 1560 AGATGTGCAGCTTGAAGAGGT-----TG 1583
Qy 820 sAspSerAspPheSerLeu-----HisAlaAlaAsnValLeuCysAr 834
Db 1584 CAATATGACTACATCTCGTTTGTATGGCCCTGAAACAATTTCTCTCATTTGCTCG 1643
Qy 834 gGluLeuAsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGly-- 853
Db 1644 G-----GTTTGTGCTGGTTCATGATCTTT 1670
Qy 854 -----AsnGlyLeuThrTrpAlaGl 860
Db 1671 CACCTCAACCCAGAACTTTCATGCTGTAGTCTTATCAGCATGGCAGTGTCCAGGAG 1730
Qy 860 uLysPheGlnCysGlu-----GlySerGluThrHisLeuAlaLe 873
Db 1731 AGGTTTCCAAAGCTGACTACTCTCCACTCTATCAGCAGCAGCAACCTCTCCAAACGAC 1790
Qy 873 uCysProIleVal-----GlnHisProGluAspThrCys----- 1885
Db 1791 GTTCCCATGCTTACTGATGGTGACAACTCTCTCCGGAATACACCTGTGGAGGTTT 1850
Qy 885 eHisSerArgGluValGlyValValCysSerArgTrpThrAspValArgLeuValAsnGl 905
Db 1851 ACTGACCTACCTATGGCAGGTTTCCAGCCCATAC----- 1887
Qy 905 yLysSerGlnCysaspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCy 925
Db 1888 -----TACCTTGGAGCTAT- 1902
Qy 925 sAspThrHisTrpaspProGluaspAlaArgValLeuCysarg----- 939
Db 1903 -----CCTAACATGCCAGATGTTTGTGGAAATTTTCGTCGCCAGCAT 1946
Qy 940 -----GlnSerCysGlyThr----- 945
Db 1947 GAACCGTGTGACAGTGTCTTTCACAGATGTCAGCTTGAAGAGGTGGCAACTATGACTA 2006
Qy 946 -AlaLeuSerThrThrGlyGlyLysTyrlleGlyGluArgSerValArgValTrpGlyHi 965
Db 2007 CATCTCGGTTTGTGATGCTGAATCAATCTTCTCTCATCTGCTCGGTT----- 2058
Qy 965 sArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThrValLeuGl 985
Db 2059 -----TGTGATGGTCCATGATCTTTCACCTCAACCCAG----- 2094
Qy 985 yAlaProCysIleHisGlyAsnThrValSerVal-----IleCysThrGlySerLe 1003
Db 2095 -----AACTTCATGCTGTAGTCTTTATCAGGATGGCAGTGT 2132
Qy 1003 uThrGlnProLeuPhe-----ProCysLeuAlaAsnValSerAs 1016
Db 2133 CACGAGGAGGTTTCCAGAGTGTACTTCCACTCTTATCAGGACCCAGCAACATCC 2192
Qy 1016 pProTyTrLeuSerAlaValProGluGly-----SerAlaLeuIleCysLeuGluAspLy 1034
Db 2193 TCCACACGCTTCCCGATCATCTGGAATGATCTTCATGTGTG----- 2238
Qy 1034 sArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyHisAs 1054
Db 2239 -----CTGAGGCTGTAATGAACAAACCGGTGTGAGGGCGAGTGAGATCTTGTACAG 2294
Qy 1054 pGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCy 1074
Db 2295 AGGCTCTTGGGTACCGTGTGCCGACGAGCTGGGACATCAATGATGCCAATGGTGTG 2354

Qy 1074 sGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGl 1094
Db 2355 CAGACAGCTCGTGTGGCTGCTCTCTGCTGCCAGAAATGCTTGGTTGGTCAGGG 2414
Qy 1094 ySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGl 1114
Db 2415 TTCAGGCTCATCTTCCTCGATGATGTCTTGTCTCTGGTATGAGTCCACCTGTGAA 2474
Qy 1114 nCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyVal 1134
Db 2475 TTGTCGTACCCCTGGCTGGCTTTCATATTTGCTGTCATGTTGAGGATGAGGAGTCAT 2534
Qy 1134 eCysSer-----GluPheThrAlaLeuArgLeuTrpSerGluThrGluTrpGluSerCy 1152
Db 2535 TTGCTCACTCCCTCATCCGACTCC-CTCTCCTGGTCCAGTTTGGACAAGTCTCTCTTTTG 2593
Qy 1152 sAlaGlyArgLeuGluValPheTyAsnGlyThrTrpGlySerValGlyArgArgAsnIl 1172
Db 2594 T-----AACTATACTTGTGGAGTTTCTCTGACTGGAC-TCTCTGGCAATTTTCTAGC 2646
Qy 1172 eThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValVa 1192
Db 2647 CCATCACTACCTGGGAGCTATCTTAATAATGCCAGATTTTGTGGA----- 2692
Qy 1192 lSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAsp----- 1209
Db 2693 -----ACATTGAAGTCCCAACACTACCGCTGACTGTGGTCTTCAGAGATGTGCGAG 2745
Qy 1210 -----IleGlnCysProLysThr-----Hl 1216
Db 2746 CTGGAAGGGGCTGCAACTATGACTATATAGAGATTTTGTATGGCCGCCACACAGTTCA 2805
Qy 1216 sIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerPro 1234
Db 2806 CCTCTCATTTGCCGGTTTGTATGGGCCCATGGCTCTTTCACCTTCAACATCCA 2860

RESULT 8

US-09-147-947-5
; Sequence 5, Application US/09147947A
; Patent No. US20020160490A1
; GENERAL INFORMATION:
; APPLICANT: TSURUOKA, No. US20020160490A1uo
; APPLICANT: YAMASHIRO, Kyoto
; APPLICANT: YAMAGUCHI, No. US20020160490A1om1
; TITLE OF INVENTION: No. US20020160490A1el Serine Protease
; FILE REFERENCE: 001560-349
; CURRENT APPLICATION NUMBER: US/09/147,947A
; CURRENT FILING DATE: 1997-03-24
; EARLIER APPLICATION NUMBER: PCT/JP98/03324
; EARLIER FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: JP 9/213969
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 2562
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; OTHER INFORMATION:
US-09-147-947-5

Alignment Scores:
Pred. No.: 4,26e-74 Length: 2562
Score: 830.00 Matches: 260
Percent Similarity: 40.26% Conservative: 106
Best Local Similarity: 28.60% Mismatches: 349
Query Match: 10.20% Indels: 195
Gaps: 34

US-09-759-130B-381 (1-1453) x US-09-147-947-5 (1-2562)

Qy	132	TrpGlySerHisAsnCysTyrHisGlyGlu-----AspValGlyVal	145
Db	112	TGGGGC-----TGCCCGCGCGGAGAGCCATCGGTCTACGCGTACGAGACTTCGGCGGC	162
Qy	146	AsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAsp-ClYAsnAsnSerCysSe	165
Db	163	CCGTGT-----CTGCGGTGGCGGAGGTGCCACCCCTTCCTGGAG	201
Qy	165	rGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAspGlyTrpAs	185
Db	202	CGGTGC-----CCCGAGCGAGCTGGCTCAGCTGCGAGGACAGC-----	241
Qy	185	nLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSerSerPheIleSerSe	205
Db	242	-----GCCACAACCTTTGTGGAGCCCGGAGCGCGGCGGAGACACC	282
Qy	205	rGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAspAspIleLeuCysG1	225
Db	283	TGCTGTTCTACGGAGACGCCCGTGCAAG-----	313
Qy	225	nGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisaspCysSe	245
Db	314	-----TGGAC-----TGGGGCTACTGCCACGTGCAG	338
Qy	245	rHisAsnGluAspValIthrLeuThrCysTyrAspSerSerAspLeuGluLeuArgLeuVa	265
Db	339	ACAGGA-----TCAGTAGCACTTCG	359
Qy	265	lGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyTh	285
Db	360	TGGCGCAAAATAGCTTTGAAGGCACATGCGAAGTATATGCAAGTGGAGTTTTGGGCGAC	419
Qy	285	rValCysHisHisLysTrpAsnAsnAlaAlaAspValValCysLysGlnLeu-----	303
Db	420	TGCTGTGAGACGCCACTGGGATGATCTGTATGCATCAGTCATTTGTCCACAGCTGCAGCT	479
Qy	304	-----GlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerG1	320
Db	480	GGGAGGAAGGAANTACGAAACAAACCCGTTTCTGCGACTG-----GG	524
Qy	320	ySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAs	340
Db	525	CCTATTCCCATTTTGGAGCAATTCGTTGCGGAGGAGATCAAGAAATATACGTGCT	584
Qy	340	pCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValI1	360
Db	585	TTGTGAAAAGACATCTGCGACGGTGGGTGTGTCTCAGAAAGATGGCAGCTGCTGCAC	644
Qy	360	eCys-----SerAspGlyAlaAspLeuGlu-----LeuArgLeuAlaAspGlySerAsnAs	377
Db	645	GTGTAGCTTTTCCCATGGGCCAACGTGCCCATCATTCGCTTCTGTGGAGCAGCAGTGT	704
Qy	377	nCysSerGlyArgValGluValArgIleHisGluGlnIleTrpTrpIleCysAspGlnAs	397
Db	705	GCATGAAGCGCGGTGGAGCTCTACCATCTGCGCCAGTGGGGAACCGTTGTGTATGACCA	764
Qy	397	nTrpLysAsnGluAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPh	417
Db	765	ATGGGATGATGCGGATGCAGAGATCTATCTGCAGGACAGCTGGCGCTCATGTGCGCAT	824
Qy	417	eGlySerArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSe	437
Db	825	AGCATGCGATCAGGCATATTTTGGGAAGGTGCTGGCCCGATTATGTTGGATGAAGTACG	884
Qy	437	rCysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgTh	457
Db	885	CTGCACCTGGGAATGAGCTTTCAATTGAGCAGTGTCCAAAGAGCTCCTGGGAGAGACATAA	944
Qy	457	rCysPheArgArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuAr	477
Db	945	CTGTGGCCATAAAGAGATGCTGGATGTCTCTGTACCCCTCTACAGACATGGGCTCATCAG	1004
Qy	477	qLeuValGlyAlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTr	497

Db 1863 AGCACACTGTTTCAAGAGGTATGGCAACAGCAGCTAGCAGCTATGCTGTTAGGGTTGGAGA 1922
QY 848 pHis-----PheGlyLysGlyAsnGly-----LeuThrTr 858
Db 1923 TTATCATCTCTGGTACCAGAGAGTTTGGAGAAATTTGGAGTTCAACAGATTGTGAT 1982
QY 858 pAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValG1 878
Db 1983 TCATCGGGAGTATCGACCCGACCGCAGTATTATGACATAGCTGTTAGATTA---CA 2039
QY 878 nHisProGluAspThrCysIleHis---SerArgGluValGlyValValCysSerArgTy 897
Db 2040 AGGACCAAGAGCAATGTCGCCAGATTCACAGCAGCCATGTTTCCACGCCCTTTACCACT 2099
QY 897 rThraspValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnVa 917
Db 2100 CTGGAGAGAGCCCAAGAAACAGCATCCAAGTGT-----TA 2138
QY 917 lLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLe 937
Db 2139 CATAACAGGATGGGT-----GACACAGGAGCAGCCCTA 2171
QY 937 uCysArgGlnLeuSerCysGlyThrAlaLeuSerThrGlyGlyLysTyrlleGlyG1 957
Db 2172 TTCAGAACACTACAA---CAACGAGCATTCCCTTACTTCTTAAAGGTTTGTGAAGA 2228
QY 957 uArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsn-----GluSe 974
Db 2229 ACCTTAAAGGTCGGTTTACAGGAGGAGTCTTGTGCTGGAACCTCCATGAACACAA 2288
QY 974 rLeuLeuAspAsnCysGlnMetThrValLeuGlyAlaPro-ProCysIleHisGlyAsnT 994
Db 2289 ACAGTGGACAGCTGCCAGGAGAC---AGCGGAGGACCACCTCATGTGTGAACGCCCGG 2345
QY 994 hrValSerValIleCysThrGly 1001
Db 2346 AGAGAGCTGGGTGGTGTATGGG- 2368

RESULT 9
US-09-918-995-32438 20030073623
; Sequence 32438, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32438
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(473)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-32438

Alignment Scores:
Pred. No.: 6,13e-74 Length: 473
Score: 818.00 Matches: 143
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.05% Indels: 0
DB: 9 Gaps: 0

US-09-759-130B-381 (1-1453) x US-09-918-995-32438 (1-473)

QY 1194 LeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysPro 1213
Db 42 CTCGCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGTGGATGACATTCAGTGTCTCT 101
QY 1214 LysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSer 1233
Db 102 AAAACGCATATCTCCATATGCGAGTGTCTGTCTGCCCATGGGAGCGAAGAAATCTCCAGC 161
QY 1234 ProAlaGluGluThrTrpIleThrCysGluAspArgIleArgValArgGlyGlyAspThr 1253
Db 162 CCAGCAGAGAGACCTGGATCACAATGTAAGATAGATAGAGTCGCTGGAGGAGACACC 221
QY 1254 GluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThrValCysAspAsp 1273
Db 222 GAGTGTCTCTGGAGAGTGGAGATCTGCGACGCGAGCTCTGGGCGACAGTGTGTGATGAC 281
QY 1274 SerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeu 1293
Db 282 TCTTGGGACCTGCCGAGGCGGAAGTGTGTGTCAGCAGCTGGGCTGTGGCTCTGCTCTG 341
QY 1294 AlaAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMet 1313
Db 342 GCTGCCCTCAGGAGCGCTTCTGTTGGCCAGGAACTGGAACCATCTGTTGGATGACATG 401
QY 1314 ArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSer 1333
Db 402 CGGTGCAAGGAATGAGTCAATTTCTATGGACTGTCTACGCCAAACCTCTGGGACAGAGT 461
QY 1334 AspCysGly 1336
Db 462 GACTGTGGA 470

RESULT 10
US-09-147-947-3
; Sequence 3, Application US/09147947A
; Patent No. US20020160490A1
; GENERAL INFORMATION:
; APPLICANT: TSURUOKA, No. US20020160490A1uo
; APPLICANT: YAMASHIRO, Kyoto
; APPLICANT: YAMAGUCHI, No. US20020160490A1el Serine Protease
; TITLE OF INVENTION: No. US20020160490A1el Serine Protease
; FILE REFERENCE: 001560-349
; CURRENT APPLICATION NUMBER: US/09/147,947A
; CURRENT FILING DATE: 1997-03-24
; EARLIER FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: JP 9/213969
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2614
; TYPE: DNA
; ORGANISM: Mouse
; FEATURE:
; OTHER INFORMATION:
US-09-147-947-3

Alignment Scores:
Pred. No.: 3,99e-63 Length: 2614
Score: 722.00 Matches: 207
Percent Similarity: 41.54% Conservative: 85
Best Local Similarity: 29.45% Mismatches: 260
Query Match: 8.87% Indels: 153
DB: 9 Gaps: 25

US-09-759-130B-381 (1-1453) x US-09-147-947-3 (1-2614)

QY 10 lLeaspPheGlyArgCysCysHisGlnAsnLeuPheSerAlaValThrCysIle 29
Db 692 GTAGACTGGGGCTACTCGCATTTGTCGAAGCCCGCGGTGGCCGTCATT----- 742
QY 30 LeuLeuAsnSerCysPheLeuIleSerSerPheAsnGlyThrAspLeuGluLeuArg 49

Db	743	----	CGC	745	----	
Qy	50	LeuValAsnGlnAspGlyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrp	69			
Db	746	CTTGTGGTGGGAACAGTGGCATGAAGTCGATGGAGCTGTACACAGCTGCCAGTGG	805			
Qy	70	GlyThrValCysAspAspGlyTrpAsnThrAlaSerThrValValCysLysGlnLeu	89			
Db	806	GGGACCATCTGTGACGACCAATGGCAATCCAGACGCAGACCTCATCTGTAGCAGCTG	865			
Qy	90	-----GlyCysProPheSerPheAlaMetPheArgPheGlyGlnAlaValThrArg	106			
Db	866	GGGCTCAGTGGCATGGCCAAAGCATGGCATCAGGCACATTTTGGGAAGGA	916			
Qy	107	HisGlyLysIleTrpLeuAspValSerCysTyrglyAsnGluSerAlaLeuTrpGlu	126			
Db	917	TCGTGGCCCAATATGTTGGATGAAGTACGCTGCACCGGAAACAGCTGTCAATGTAGACAA	976			
Qy	127	CysGlnHisArgGluTrpGlySerHisAsnCysTyrrHisGlyGluAspValGlyValAsn	146			
Db	977	TGTCCAAAGAGTTCCTGGGGCGCAACATAACTGTGGCCATAAAGAGAGATGCTGGAGTGTCT	1036			
Qy	147	CysTyrglyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSergly	166			
Db	1037	TGTGTCTTCTTAACAGATGGTGTCTATCAGATGGCGAGGAGGAAAGTACCACATGAAGT	1096			
Qy	167	ArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeu	186			
Db	1097	CGCTGGAGGCTTACTCAAGGGGACGTGGGGACAGTCTGTGATGATGCTGGACTGTAG	1156			
Qy	187	AsnThrAlaAlaValValCysArgGlnLeuGlyCysProSerSerPheIleSerSergly	206			
Db	1157	ATGAACACATACGTGGCTGTCTGCATCTGTGGGATTTAAATACGCCAACAGTCTCTCT	1213			
Qy	207	ValValAsnSerProAlaValLeuArgProIleTrpLeuAspAspIleLeuCysGlnGly	226			
Db	1214	GTAAACCATTTTGTATGGCAGCAACAGGCCCATATGGCTGGATGACGTACAGTGTCTCAGGA	1273			
Qy	227	AsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHis	246			
Db	1274	AAAGAAGTCAAGTTCATTCAGTGTCCAGGAGACAGTGGGGAAGGCATGACTGCAGCCAT	1333			
Qy	247	AsnGluAspValThrLeuThrCysTyrrAspSerSerAsp	259			
Db	1334	AGAGAAGATGTGGGCTTCACCTGCTATCCTGCACGGATGGACATAGCCTTCTCCAGGT	1393			
Qy	260	LeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIle	279			
Db	1394	TTTCCCATCAGCATAGTGGATGGAGAGAATAAGAAGGAAGCAGCTGGAGGTTTGTCT	1453			
Qy	280	GlnGlyArgTrpGlyThrValCysHisIleLysTrpAsnAsnAlaAlaAaspValVal	299			
Db	1454	AATGGCCATGGGGAACAATCTCGCATGACGGATGGACCGGATAGCATGACGTGTGATC	1513			
Qy	300	CysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSer	319			
Db	1514	TGCGGCGAGCTTGGCTATGAAGGTCCTGCCAGACAAGCATATGGCTTATTTGGGGAA	1573			
Qy	320	GlySerAspValValTrpLeuAspGlyValSerCysSerglyAsnGluSerPheLeuTrp	339			
Db	1574	GGAAAAGGCCCATCCATCGATATGTGAAGTGCACAGAAATGAGAAGGCCCTGGCT	1633			
Qy	340	AspCysArgHisSerglyThrValAsnPheAspCysLeuHisGlnAsnAspValSerVal	359			
Db	1634	GACTGTGTCAACAAAGACATTTGGAAGGCACAATCTGCCCCACAGTGAAGATGCAGGAGTC	1693			
Qy	360	IleCys-----SerAspGlyAlaAspLeuGluLeuArgLeuAla	372			
Db	1694	ATCTGTGACTATTTACAGAGAAGAACCATCACTAGTGTGTATTAAGAGATGCTC	1747			
Qy	373	AspGlySerAsnAsnCysSerglyArgValGluValArgIleHisGlu-----GlnTrp	390			

D	b	1748	-----TCATCTGGAGTGGACTGAGGTTCACGCCGTCG--GCAGAAACGGATCATTTGG	1800
Q	y	391	TrpThrIleCysAspGlnAsnTrpLysAsnGluAlaLeuValValCysLysGlnLeu	410
D	b	1801	TGG-----GAACAATTCTTAAGGGTGCCCTGGCCTTGCCA-----	1836
Q	y	411	GlyCysProPheSerValPheGlySerArgArgAlaLysProSerAsnGluAlaArgAsp	430
D	b	1837	GGCTTCCCTCAGGCTGAGGCTCGGCCCTGGAGACGCGCAGGCTGCT-----	1881
Q	y	431	IleTrp-----lleAsnSerIleSerCysThrGlyAsnGlu-----	442
D	b	1882	TTGTGGAGCTACCCTTCTGAGTAGCTGTGGGTCTCTGACAGCTGCACATGCTTCAAAAG	1941
Q	y	443	-----SerAlaLeuTrpAspCysCysThrTyrrAspGly-----	452
D	b	1942	GTACGGAAACAACCTCGAGGAGCTATCCAGTTCGAGTTGGGATTATCATCTCTGTCTACC	2001
Q	y	453	-----LysAlaLysArgThrCysPheArgArgSerAspAlaGly	465
D	b	2002	AGAGGAGTTGAACAAGAANAATAGGGGTTCAACAGATTGT-----GATTCCACAG	2049
Q	y	466	ValIleCysSer-AspLysAlaAspLeuAspLeuArgLeuValGlyAlaHisserProCy	485
D	b	2050	GAACATACAGGCCACAGAACGAGCATGACATGACATGCCCTGGTTAGATTGCAAGGACCA--	2107
Q	y	485	stfyrGlyArgLeuGluValLysTyrrGlnGlyGluTrpGlyThrValCysHisAspargtr	505
D	b	2108	----GGGGACAATGCCAGACTAAGCACCCACGCTTTTGGCCAGCCTGTTTACCTCTATG	2163
Q	y	505	pSerThrArg-----AsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysProm	523
D	b	2164	GAGAGAGAGGCCACAGAAACACGCCCTCCAACTGTCACATAACAGGATGGGGAGACACAGG	2223
Q	y	523	thisValPheGlyMetThr-----TyrPh	531
D	b	2224	TGTCGCTTACTCAAGNACTCTACACAAGCTGCTGTGCCTCTGTTTACCCAGAGGTTTTG	2283
Q	y	531	elysGluAlaSerGlyProIleTrpLeuAspValSerCysIleGlyAsnGluSerAs	551
D	b	2284	TAAAGAGAGGTACAAGGGGACTATTACTGGGAGAAATGCTCTGTCTGTCTGGGAAC-----	2335
Q	y	551	nleirpasCysGluHisSerGlyTrpGlyLysHisAsnCysValHisargGluAspVa	571
D	b	2336	-----CTCCAAGAACACACCG	2352
Q	y	571	IleValThrCysSerGlyAspAlathrtpGlyLeuArgLeuValGlyGlySerAsnAr	591
D	b	2353	TGTGGACAGCTGCCGGGAGACAGT-----GGAGGACCACTCAT	2391
Q	y	591	gCysSerGlyArgLeuGlu-----ValTyrPheGlnGlyArgTrpGlyThrValCy	608
D	b	2392	GTGTGAAAACGCTGATGAGTCTCGGTTGTGTATGGGTGACTTCCTCGGGGGTATGGATG	2451
Q	y	608	s-----AspaspGlyTrpAsnSerLysAlaAlaValValCysSercInLeuas	625
D	b	2452	TGGAGTCAAAGACAC-TCTCTGGAGTTTATACCAGAGTCCCCGCTTTGTACC---TTGGA	2507
Q	y	625	p-----CysProSerIleilecglyMetGlyLeuGlyAsnAlaSerThrGlyTyrrGl	643
D	b	2508	TAAAAAGTCTCACCAGTCTGTACT-----TATGG	2537
Q	y	643	yLysIle	645
D	b	2538	AAAGCTC	2544

RESULT 11
US-09-905-291A-147
; Sequence 147, Application US/09905291A
; Patent No. US2002016037A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi

Qy 345 GlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer 362
Db 1082 TTTTGGGGTTCACGACTGCACCCAGGAGATGTGGTGTCTCATCTGCTCA 1135

RESULT 12

US-09-902-853-147

; Sequence 147, Application US/09902853

; Publication No. US20020192659A1

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth, J.

; APPLICANT: Kijavlin, Ivar J.

; APPLICANT: Mather, Jennie P.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas P.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/902,853

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: US/09/665,350

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/21547

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/23089

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: PCT/US99/28214

; PRIOR FILING DATE: 1999-11-29

; PRIOR APPLICATION NUMBER: PCT/US99/28313

; PRIOR FILING DATE: 1999-11-30

; PRIOR APPLICATION NUMBER: PCT/US99/28564

; PRIOR FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: PCT/US99/28565

; PRIOR FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: PCT/US99/30095

; PRIOR FILING DATE: 1999-12-16

; PRIOR APPLICATION NUMBER: PCT/US99/30911

; PRIOR FILING DATE: 1999-12-20

; PRIOR APPLICATION NUMBER: PCT/US99/30999

; PRIOR FILING DATE: 1999-12-20

; PRIOR APPLICATION NUMBER: PCT/US00/00219

; PRIOR FILING DATE: 1999-12-20

; PRIOR APPLICATION NUMBER: PCT/US00/00219

; PRIOR FILING DATE: 1999-12-20

; PRIOR APPLICATION NUMBER: PCT/US00/00219

; PRIOR FILING DATE: 1999-12-20

; PRIOR APPLICATION NUMBER: PCT/US00/00219

; PRIOR FILING DATE: 1999-12-20

; PRIOR APPLICATION NUMBER: PCT/US00/00219

; PRIOR FILING DATE: 1999-12-20

; PRIOR FILING DATE: 2000-01-05

; NUMBER OF SEQ ID NOS: 423

; SEQ ID NO 147

; LENGTH: 1686

; TYPE: DNA

; ORGANISM: Homo Sapien

US-09-902-853-147

Alignment Scores:

Pred. No.: 2,64e-60 Length: 1686

Score: 691.50 Matches: 148

Percent Similarity: 52.51% Conservative: 40

Best Local Similarity: 41.34% Mismatches: 139

Query Match: 8.50% Indels: 31

DB: 9 Gaps: 9

US-09-759-130b-381 (1-1453) x US-09-902-853-147 (1-1686)

Qy 21 LeuPheSerAlaValThrCysIleLeuLeuLeuAsnSerCys-----Phe 36
Db 107 CTATTCTCC-----TTGATCCTTGGCCATTTCACCCAGACCTGGATTTC 148
Qy 37 LeuIleSerSerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyPro 56
Db 149 CTAGCCTCTCCATCTGGA-----GTGGCGCTGGTGGGGGCTCCACCGC 193
Qy 57 CysSerGlyThrValGluValLysPheGlnGlyIntrpGlyThrValCysAspGly 76
Db 194 TGTGAAGGGCGGTGGAGGTGGACAGAAAGCCAGTGGGGCAGCGTGTGTGATGACGCG 253
Qy 77 TrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAla 96
Db 254 TGGGACATTAAAGGACGTGGCTGTGTGGCGGGAGCTGGGCTGTGGAGCTGCCAGCGGA 313
Qy 97 Met-----PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAsp 113
Db 314 ACCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAGAACCAAGGTCTCATCCAA 373
Qy 114 AspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGly 133
Db 374 TCAGTCAAGTTCACACAGGAACAGAGATACATTGGCTTCAGTGTGACGAAGAAGAA----- 427
Qy 134 SerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys----- 147
Db 428 GTTTATGATTTTCACATGATGAAGATGCTGGGGCATGCTGTGAGAACCCAGAGACTCT 487
Qy 148 TyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArg 167
Db 488 TTCTCCCGCAGTCCAGAGGGTGTGAGGCTGGCTGAGCGGCCCTGGGCATTCAGAGGACGC 547
Qy 168 ValGluValLysPheGlnGluArgTrpGlyThrIleCysAspGlyTrpAsnLeuAsn 187
Db 548 GTGGAAGTGAAGCACCAGCAGTGGTATACCGTGTGCCAGACAGGTGGAGCTCCGG 607
Qy 188 ThrAlaAlaValValCysArgGlnLeuGlyCysProSerSerPheIleSerGlyVal 207
Db 608 GCCGCAAGGTGTGTGCCGCGAGCTGGAGTGTGGAGGGCTGTACTGACTCAAAACGC 567
Qy 208 ValAsnSerProAlaValLeuArg---ProIleTrpLeuAspAspIleLeuCysGlnGly 226
Db 668 TGCAACAAGCATGCCTATATGGCCGAAAAACCATCTGGCTGAGCCAGATGTCATGCTCAGGA 727
Qy 227 AsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisCysSerHis 246
Db 728 CGAAGCAACCCCTTCAGAGTTCCTTTGGGGCCTTTGGGGGGAAGAACACCTGCACACCT 787
Qy 247 AsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuLeuArgLeuValGly 266
Db 788 GATGAAGACACGTGGGTGCAATGTGAAGATCCCTTTGAC-----TTGAGACTAGTAGGA 841
Qy 267 GlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyThrVal 286
Db 842 GGAGACAACCTCTGCTCTGGCGGAGTGGAGGTGCTCCACAGGCGCTATGGGCTCTGTC 901

```

: PRIOR FILING DATE: 1999-11-29
: PRIOR APPLICATION NUMBER: PCT/US99/28313
: PRIOR FILING DATE: 1999-11-30
: PRIOR APPLICATION NUMBER: PCT/US99/28564
: PRIOR FILING DATE: 1999-12-02
: PRIOR APPLICATION NUMBER: PCT/US99/28565
: PRIOR FILING DATE: 1999-12-02
: PRIOR APPLICATION NUMBER: PCT/US99/30095
: PRIOR FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: PCT/US99/30911
: PRIOR FILING DATE: 1999-12-20
: PRIOR APPLICATION NUMBER: PCT/US99/30999
: PRIOR FILING DATE: 1999-12-20
: PRIOR APPLICATION NUMBER: PCT/US00/00219
: PRIOR FILING DATE: 2000-01-05
: NUMBER OF SEQ ID NOS: 423
: SEQ ID NO 147
: LENGTH: 1686
: TYPE: DNA
: ORGANISM: Homo Sapien
US-09-907-824-147

Alignment Scores:
Pred. No.: 2,64e-60 Length: 1686
Score: 691.50 Matches: 148
Percent Similarity: 52.51% Conservative: 40
Best Local Similarity: 41.34% Mismatches: 139
Query Match: 8.50% Indels: 31
DB: Gaps: 9

US-09-759-130B-381 (1-1453) x US-09-907-824-147 (1-1686)
Qy 21 LeuPheSerAlaValThrCysIleLeuLeuLeuAsnSerCys-----Phe 36
    |||||
Db 107 CTATTCCTCC-----TTGATCCTTCCCATTTGACACACCTGGATTC 148

Qy 37 LeuIleSerSerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyPro 56
    ||| |||
Db 149 CTACGGCTTCATCTGGA-----GTGCGGCTGTGGGGGCTCCACCGC 193

Qy 57 CysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGly 76
    |||||
Db 194 TGTGAAGGGGGTGGAGGTGGAAACAGAAAGCCAGTGGGGCACCGTGTGTGATGACGC 253

Qy 77 TrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAla 96
    ||:|:|:|
Db 254 TGGGACATTAAGGACCTGGCTGTGTCTGCGGGAGCTGGGCTGTGGAGCTGCCACGGA 313

Qy 97 Met-----PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAsp 113
    :|:|:|
Db 314 ACCCCTAGTGTATTTTGTATGAGCCACACGAGCAAAAGAGCAAAAGGTCCTCATCCAA 373

Qy 114 AspValSerCysTyrGlyAsnGluSerAlaLeuTrpGlyCysGlnHisArgGluTrpGly 133
    |||||
Db 374 TCAGTCAAGTTCACAGACAGAACAGATACATTCGCTGAGTGTGAGCAAGAAGAA----- 427

Qy 134 SerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys----- 147
    :|:|:|:|
Db 428 GTTTATGATGTTTCATGATGATGATGCTGGGCGCATCTGTGTGAGAACCCAGACAGCTCT 487

Qy 148 TyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArg 167
    :|:|
Db 488 TTCTCCCGAGTCCACAGAGGGTGTGAGCTGGCTGACGCGCCTGGGCATTCGAAGGACGC 547

Qy 168 ValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsn 187
    |||||
Db 548 GTGGAAGTGAAGCACACAGAACCATGCTGATACCGTGTGTCGAGACAGAGCTGGAGCCCTCGG 607

Qy 188 ThrAlaAlaValValCysArgGlnLeuGlyCysProSerSerPheIleSerSerGlyVal 207
    |||
Db 608 GCCCGAAGGTGGTGTGCCGCGAGCTGGGATGTGGGAGGCGCTGTACTGACTCAAAACGC 667

Qy 208 ValAsnSerProAlaValLeuArg---ProIleTrpLeuAspAspIleLeuCysGlnGly 226

```



```

; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423

```

```

: SEQ ID NO 147
:
: LENGTH: 1686

```

```

: TYPE: DNA
: ORGANISM: Homo Sapien
US-09-904-011-147

Alignment Scores:
Pred. No.: 2,64e-60 Length: 1686
Score: 691.50 Matches: 148
Percent Similarity: 52.51% Conservative: 40
Best Local Similarity: 41.34% Mismatches: 139
Query Match: 8.50% Indels: 31
DB: 9 Gaps: 9

US-09-759-1308-381 (1-1453) x US-09-904-011-147 (1-1686)

Qy 21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCys-----Phe 36
|||||
Db 107 CTATCTCC-----TTGATCCTGCCATTGACACAGACCTGGATTC 148
|||||

Qy 37 LeuLeSerSerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyPro 56
|||||
Db 149 CTACGGCTCTCATCTGGA-----GTGCGGCTGGTGGGGGGCCTCCACCGC 193
|||||

Qy 57 CysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspGly 76
|||||
Db 194 TGTGAAGCGCGGTGGAGGTGGACAGAAAGCCAGTGGGGCACCTGTGTGATGACGCG 253
|||||

Qy 77 TrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAla 96
|||||
Db 254 TGGGACATTAAGGACGTGGCTGTGTGTCGCGGAGCTGGGCTGTGGAGCTGCCAGCGGA 313
|||||

Qy 97 Met-----PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAsp 113
|||||
Db 314 ACCCCTAGTGGTATTTGTTAGGCCACCACAGCAAAAGAGCAAAAGGTCTCATCCAA 373
|||||

Qy 114 AspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGly 133
|||||
Db 374 TCAGTCAGTTCGCAGGACAGAGATACATTCGCTCAGTGTGAGCAAGAGAA----- 427
|||||

Qy 134 SerHisAsnTyrHisGlyGluAspValGlyValAsnCys----- 147
|||||
Db 428 GTTTATGATTGTTACATGATGAAGATGCTGGGGGCATCGTGTGAGAACCCAGAGAGCTCT 487
|||||

```

148	TyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArg	167
Qy	:::	
488	TTCTCCCCAGTCCCAAGAGGGTGTCCAGGCTGGCTGACGGCCTGGCATTCGAGGACGC	547
Db	:::	
168	ValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsn	187
Qy		
548	GTGGAGTGAACACCAACAGTGGTATACCGTGTGCCAGACAGCTGGACGCTCCCG	607
Db		
188	ThrAlaAlaValValCysArgGlnLeuGlyCysProSerSerPheIleSerGlyVal	207
Qy		
608	GCGCACAAAGTGGTGTCCGCGACGTGGATGTGGAGGGCTGTACTGACTCAAAACGC	667
Db		
208	ValAsnSerProAlaValLeuArg---ProIleTrpLeuAspAlleLeuCysGlnGly	226
Qy		
668	TGCAACAAGCATGCCATATGGCGCGAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGA	727
Db		
227	AsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHis	246
Qy		
728	CGAAGAAGCAACCTTCAGAGATTGCCCTTCTGGGCTTGGGGGAAGAACACCTGCAACCAT	787
Db		
247	AsnGluAspValThrLeuThrCysTrpAspSerSerAspLeuGluLeuArgLeuValGly	266
Qy		
788	GATGAAGACACCTGGTGCATGTGAAGATCCCTTTGAC-----TTGACACTAGTAGGA	841
Db	:::	
267	GlyThrAsnArgCysMetClyArgValGluLeuLysIleGlnGlyArgTrpGlyThrVal	286
Qy		
842	GGAGACAACTCTGCTTGGGCGACTGGAGGTGTGCAACAGGCGTATGGGCTCTGTC	901
Db		
287	CysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCysLysGlnLeuGlyCysGly	306
Qy		
902	TGTTGATCAACTGGGGAGAAAAGGAGGACCAAGTGTATGCAAGCAACTGGCTGTGG	961
Db		
307	ThrAlaLeu-----HisPheAlaGlyLeuProHisLeuGlnSerGlySerAspValVal	324
Qy		
962	AAGTCCCTCTCCCTCCCTTCAGAGACCGGAATGCTATGGCCCTGGGGTGTGCCGCATC	1021
Db		
325	TrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCysArgHisSer	344
Qy		
1022	TGGCTGGATATGTTCTGTTGCTCAGGGGAGGAGCAGTCCCTGGAGCACTGCCAGCACAGA	1081
Db		
345	GlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer	362
Qy		
1082	TTTTTGGGGGTTTCAGCATGTGACCAACCAAGCAAGATGTGGCTGTCACTGCTCA	1135
Db		

Search completed: May 12, 2003, 17:37:41
Job time : 376 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 12, 2003, 06:37:56 ; Search time 446 Seconds
(without alignments)
7336.667 Million cell updates/sec

Title: US-09-759-130B-381

Perfect score: 8138

Sequence: 1 MMLPQNSWIDFGRCCHQN.....CEDASDTSLGLVPASEATK 1453

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US0759130/runat_06052003_123553_23343/app_query.fasta.1.1607
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -NORM=ext -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0759130.ecgn.1.1.301 @runat_06052003_123553_23343 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	8138	100.0	4359	22	Human TANGO 234 OR
2	8138	100.0	4628	22	Human TANGO 234 CD
3	6107.5	75.0	3670	22	DNA encoding human
4	4712	57.9	2800	22	Human full-length
5	3429	42.1	3716	22	Human polynucleoti
6	3429	42.1	3834	22	Human EST-derived
7	3429	42.1	3834	22	Human polynucleoti
8	3429	42.1	3834	22	Human polynucleoti
9	3396.5	41.7	3811	22	Human polynucleoti
10	3377	41.5	3707	22	Human DNA sequence
11	3082	37.9	4308	22	Bovine WCL ORF. B
12	2435	29.9	5802	19	Human SRCR protein
13	1763.5	21.7	2697	24	DNA encoding novel
14	1404	17.3	820	21	Human secreted exp
15	1183	14.5	690	22	Human protein enco
16	1142.5	14.0	4344	24	Rat sequence diffe
17	1142.5	14.0	4360	18	Rat von Ebner's gl
18	1081	13.3	608	20	EST clone Crl18.
19	1033	12.7	562	22	Human protein enco
20	972.5	12.0	1659	22	Human sbg14862SPER
21	921	11.3	1269	22	Human sbg14862SPER
22	830	10.2	2562	20	Human serine prote
23	830	10.2	3350	20	Human neurotropsin
24	785.5	9.7	29598	19	Human Sc2 DNA. Ho
25	727	8.9	450	22	Mouse neurotropsin
26	722	8.9	2376	20	Mouse serine prote
27	722	8.9	2614	20	Human SRCR protein
28	712.5	8.8	2000	19	Human SRCR protein
29	691.5	8.5	1686	20	Protein PRO229 CDN
30	691.5	8.5	1686	22	Human PRO229 CDNA.
31	691.5	8.5	1686	24	Human angiogenesis
32	691.5	8.5	1686	24	Human PRO229 CDNA
33	691.5	8.5	1686	24	CDNA encoding huma
34	691.5	8.5	2181	19	Human scavenger re
35	690.5	8.5	1041	19	Human liver cell c
36	690.5	8.5	1591	19	Human liver cell c
37	684	8.4	527	22	Human contig polyn
38	647.5	8.0	3767	21	Polynucleotide iso
39	639.5	7.9	3532	24	Human C6153 (or C5
40	638	7.8	2420	22	Human C6153 (or C5
41	635.5	7.8	5059	22	Human C6153 (or C5
42	634.5	7.8	2262	21	Human lysyl oxidas
43	634.5	7.8	2262	22	Nucleotide sequenc
44	634.5	7.8	2785	24	Human secreted pro
45	634.5	7.8	2920	21	Human lysyl oxidas

ALIGNMENTS

RESULT 1

AAF45124

ID AAF45124 standard; CDNA; 4359 BP.

AC AAF45124;

DT 30-MAR-2001 (first entry)

DE Human TANGO 234 ORF.

XX

XX

XX

XX

XX

XX

XX

XX

XX

OS

Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
central nervous system; focal brain disorder; bipolar affective disease;
global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
neuropsychiatric; psychoactive substance use; anxiety; ss.
Homo sapiens.

Db 1441 GCTCATAGCCCCCTGTATGGAGATGGAGGTGAAATACCAAGGAGAGTGGGGGACTGTG 1500
Qy 501 CysHisAspArgTrpSerThrArgAsnAlaAlaValCysLysGlnLeuGlyCysGly 520
Db 1501 TGTCTATGACAGATGGAGCACAGGAATGACGTGTGTGTGTAAACAATTTGGGATGTGGA 1560
Qy 521 LysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeu 540
Db 1561 AAGCCTATGCATGTCTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTTGGCTG 1620
Qy 541 AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp 560
Db 1621 GATGACGTTTCTGCAATTTGGAAATGAGTCAAAATATCTCGGACTGTGAACACAGTGGATGG 1680
Qy 561 GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr 580
Db 1681 GGAAGCATATTTGTACACAGAGAGATGTGATGTAACTGCTCAGGTGATGCCACA 1740
Qy 581 TrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPhe 600
Db 1741 TGGGGCTGAGGCTGGTGGGGGAGCAACCGCTCTCGGGAAGACTGGAGGTGTACTTT 1800
Qy 601 GlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValVal 620
Db 1801 CAAGGACGCTGGGGCACAGTGTGTATGACGCTGGAACAGTAAAGCTGCAGCTGTGGTG 1860
Qy 621 CysSerGlnLeuAspCysProSerSerIleGlyMetGlyLeuGlyAsnAlaSerThr 640
Db 1861 TGTACCCAGCTGGACTGCCCATCTTCTATCATTTGGCATGGGTCTGGGAACGCTTCTACA 1920
Qy 641 GlyTyrGlyLysIleTrpLeuAspValSerCysAspGlyAspGluSerAspLeuTrp 660
Db 1921 CGATATGGAATAATTTGGCTCGCATGCTTCTCTGTATGGAGATGAGTCAAGTCTCTGG 1980
Qy 661 SerCysArgAsnSerGlyTrpClyAsnAsnAspCysSerHisSerGluAspValGlyVal 680
Db 1981 TCATCAGGAACAGTGGGTGGGGAATAATGACTGCAGCTCACAGTGAAGATTTGGAGTG 2040
Qy 681 IleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyCysSerArgCysAla 700
Db 2041 ATCTGTTCTGATGCATCGGATATGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCT 2100
Qy 701 GlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGly 720
Db 2101 GGAAGAAGTTGAGTGAATGCTCCAGGCTGCCGTGGGAATCTGTGCTAAATGCCGTGGGA 2160
Qy 721 MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer 740
Db 2161 ATGAACATTTCTGGAAGTTGTTTGCAGGCAACTGTAATGTGGGTCTGCAATCAGGCTCTCC 2220
Qy 741 ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr 760
Db 2221 ACAGAGCCTCATTTTCACAGAAAGACATTTACACATCTTAATGTCGAATTTCTGGCTGCACT 2280
Qy 761 GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis 780
Db 2281 CGAGGGGAAGCCTCTCTCTGGGATGTATACGATGGGAGTGGAAACACAGACTCGGTGTAT 2340
Qy 781 LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAla 800
Db 2341 TTAATAATGGAAGCAAGTTTGTATCTGCAGCCCCACAGCCAGCCGCTGGTGGAGCT 2400
Qy 801 AspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys 820
Db 2401 GATATGCCCTGCTCTGGAGCTGTTGAGTGAACATGACACACATGCGCTCTGTCTGT 2460
Qy 821 AspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp 840
Db 2461 GATTCTGATTTCTCTCTCATGCTGCCAATGTGCTGTCAGAGAAATTAATTTGTGGAGAT 2520
Qy 841 AlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu 860
Db 2521 GCCATATCTCTTCTGTGGGAGATCATTGGAAGGGAATGGTCTAACTTTGGGCCGAA 2580

Qy 861 LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro 880
Db 2581 AAGTTCAGTGTGAAGGAGTGAACCTACCTGCTGATATATGCCCCATTTGTTCAACATCG 2640
Qy 881 GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspVal 900
Db 2641 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCGCGATATACAGATGTC 2700
Qy 901 ArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHis 920
Db 2701 CGACTTGTGAATGGCAATCCAGTGTGACGGCAAGTGGAGATCAACGTCTTGACAC 2760
Qy 921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940
Db 2761 TGGGGCTCAGTGTGTGACACCCACTGGGACCCAGAGATGCCCGTCTCTATGACAGACAG 2820
Qy 941 LeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerVal 960
Db 2821 CTCAGCTGTGGGACTGCTCTCAACACAGAGGAAAAATATATTGGAAAAAGATGTT 2880
Qy 961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGln 980
Db 2881 CGTGTGTGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGATTAACGTCAA 2940
Qy 981 MetThrValLeuGlyAlaProCysIleHisGlyAsnThrValSerValIleCysThr 1000
Db 2941 ATGACAGTCTTGGAGCACCTCCCTGTATCCATGAAATACTGCTCTCTGATCTGCACA 3000
Qy 1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSer 1020
Db 3001 GGAAGCCTGACCCAGCCACTGTTCCATGCCCTCGAAATGTATCTGACCCATATTGTCT 3060
Qy 1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuLeuValAsp 1040
Db 3061 GCAGTTCACAGAGGAGTCTTGTATCTCTTAGAGGACAAAGGCTCCGCCCTAGTGGAT 3120
Qy 1041 GlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIle 1060
Db 3121 GGGGACAGCCGTGTGCCGGAGAGTAGAGATCTATCACAGCGCTTCTGGGCGACCATC 3180
Qy 1061 CysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080
Db 3181 TGTGATGAGCGGTGGGACCTGAGCGATGCCACGTGGTGTGTCAAAGTGGGCTGTGGA 3240
Qy 1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeu 1100
Db 3241 GTGGCCTTCAATGCCACGGTCTCTGCTCACTTTGGGAGGGGTGAGGCGCCATCTGGCTG 3300
Qy 1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120
Db 3301 GATGACCTCAACTGCACAGGAACGGAGTCCCCTTGTGGCAGTGGCTTCCCGCGCTGG 3360
Qy 1121 GlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAla 1140
Db 3361 GGGGACAGCAGCTGAGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3420
Qy 1141 LeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyr 1160
Db 3421 TTGAGGCTCTACAGTGAACACTGAACAGAGAGCTGTGCTGGGAGATTTGAAGTCTTCTAT 3480
Qy 1161 AsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleVal 1180
Db 3481 AACGGACCTGGGCGAGCGTCGGCAGGAGGAACATCACACAGCCATACGAGCATTTG 3540
Qy 1181 CysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThr 1200
Db 3541 TGCAGGACAGCTGGGCTGTGGGAGAAATGGAGTTGTGAGGCTCGGCCCTTTATCTAAGACA 3600
Qy 1201 GlySerGlyPheMetTrpValAspAlaGlnCysProLysThrHisIleSerIleTrp 1220
Db 3601 GGCTCTGTTTCTATGCTGGTGGATGACATTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCT 3660

QY 1221 GlnCysLeuSerAlaProTrpGluArgIleSerSerProAlaGluThrTrpIle 1240
 DB 3661 CAGTGCCTCTGCCCATGGGAGGAGAAATCTCAGCCAGAGAGACCTGGATC 3720
 QY 1241 ThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyValGlu 1260
 DB 3721 ACATGTGAAGATGAATAGAGTGGTGGAGGAGACCGAGTCTCTGGAGAGTGGAG 3780
 QY 1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspSerTrpAspLeuAlaGluAla 1280
 DB 3781 ATCTGGCAGCGAGCTCTGGGGCAGAGTGTGTGATGACTCTGGACCTGCCGAGGG 3840
 QY 1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300
 DB 3841 GAAGTGGTGTGTCAGCAGCTGGCTGTGGCTCTGTCTGTGGCTGCCCTGAGGACGCTTCG 3900
 QY 1301 PheGlyGlnGlyThrClyThrIleTrpLeuAspMetArgCysGlyAsnGluSer 1320
 DB 3901 TTTGGCCAGGGAACCTGAACCATCTGGTGGATGACATCGGTGCAAGAGAAATGAGTCA 3960
 QY 1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340
 DB 3961 TTTCTATGGAGCTGTACACCCAAACCTGGGACAGAGTGTGGACACAGGAGAT 4020
 QY 1341 AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeu 1360
 DB 4021 GTGGCGTGAAGTGTCTGGCAGTGTCTGAAATCACTGAATGCCTCTCAGGTCAATTA 4080
 QY 1361 AlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPheIleLeuPheLeu 1380
 DB 4081 GCATTTATTTATCCAGTATCTTTGGGCTCTCTCTGGTGTCTGTTATTCTATTCTC 4140
 QY 1381 ThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArg 1400
 DB 4141 ACCTGTGTCGAGTTCAGAAACAAACATCTGCCCTCAGAGTTTCAACCAAGAGGAGG 4200
 QY 1401 GlySerLeuGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspPro 1420
 DB 4201 GGTTCCTCTCGAGGAGAAATTTATTCATGAGATGGAGACCTGCTCAAGAGAGGACCCA 4260
 QY 1421 HisGlyThrArgThrSerAspThrProAsnHisGlyCysGluAspAlaSerAspThr 1440
 DB 4261 CATGGACAAAGACCTCAGATGACACCCCAACCATGGTGTGTGAAGATGCTAGCGACACA 4320
 QY 1441 SerLeuLeuGlyValLeuProLaserGluAlaThrLys 1453
 DB 4321 TCGCTGTGGAGTCTCTCTGCTGCTGAAGCCACAAAA 4359
 RESULT 2
 AAF45123
 ID AAF45123 standard; cDNA; 4628 BP.
 XX
 AC AAF45123;
 DT 30-MAR-2001 (first entry)
 XX
 DE Human TANGO 234 cDNA.
 XX
 KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 KW neuropsychiatric; psychoactive substance use; anxiety; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200077239-A2.
 XX
 PD 21-DEC-2000.
 XX
 PF 24-MAY-2000; 2000WO-US14858.
 XX

PR 14-JUN-1999; 99US-0333159.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
 XX
 XX WPI; 2001-032313/04.
 DR P-PSDB; AAB66037.
 XX
 XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -
 XX
 PS Claim 1; Fig 2; 359pp; English.
 XX
 CC The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.
 XX
 SQ Sequence 4628 BP; 1174 A; 925 C; 1352 G; 1177 T; 0 other;
 Alignment Scores:
 Pred. No.: 0 Length: 4628
 Score: 8138.00 Matches: 1453
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0
 US-09-759-130B-381 (1-1453) x AAF45123 (1-4628)
 QY 1 MetMetLeuProGlnAsnSerTrpHisIleAspPheGlyArgCysCysHisGlnAsn 20
 DB 28 ATGATGCTGCCCTCCAAACCTCGGGCATATGATTTGGAGATGCTGCTGTCATCAGAAC 87
 QY 21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSerSer 40
 DB 88 CTTTCTCTGCTGTGGTAACTTGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 147
 QY 41 PheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr 60
 DB 148 TTTAATGGAACAGATTGGAGTTGAGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 207
 QY 61 ValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThr 80
 DB 208 GTGGAGGTGAATTCACGAGCAGTGGGGGACTGTGTGTGTGTGTGTGTGTGTGTGTGT 267
 QY 81 AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArgPhe 100
 DB 268 GCCTCAACTGCTCGTGGCAACAGCTTGGATGTCCATTTTCTTCGCCATGTTTCGTTTT 327
 QY 101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsn 120
 DB 328 GGACAAAGCCGTGACTAGACATGAAAAATTTGGCTTGTGTGTGTGTGTGTGTGTGTGT 387
 QY 121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140
 DB 388 GAGTCAGCTCTCTGGGAATGTCAACACCGGAATGGGAGGCCATTAACCTGTATCATGGA 447
 QY 141 GluAspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGly 160
 DB 448 GAAGATGTTGGTGTGAACCTGTTATGTGTGAAGCAATCTGGGTTTGAGGCTAGTGGATGGA 507

Db	1588	AGCCATGATGCTGTTGGTATGACCTATTTAAAGAAGCATCAGGACCTATTGGCGT	164
Qy	541	AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp	560
Db	1648	GATGACCTTCTTGGCATGGAAATGAGTCAAAATATCTGGACATGTCACACAGTGGATGG	1707
Qy	561	GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr	580
Db	1708	GGAAAGCATAAATTGCTGTACACAGAGAGGATGTGATTGTAACTGCTCAGGTGATGCAACA	1767
Qy	581	TrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTrpPhe	600
Db	1768	TGGGGCTGAGGCTGGTGGCGGCACACCGCTGCTCGGGAAGACTGGAGGTGTACTTT	1827
Qy	601	GlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValVal	620
Db	1828	CAAGGACGGTGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTG	1887
Qy	621	CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr	640
Db	1888	TGTAGCCAGCTGGACTGCCCATCTTCTATCATTTGGCATGGGTGCGGAACGCTTCTACA	1947
Qy	641	GlyTrpGlyLysIleTrpLeuAspValSerCysAspGlyAspGluSerAspLeuTrp	660
Db	1948	GGATATGGAAAAATTTGGCTCGATGATGTTTCTCTGTGTGGAGATGAGTCAGATCTCTGG	2007
Qy	661	SerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyVal	680
Db	2008	TCATGTCAGGAACAGTGGTGGGAAATAATGACTGCAGTCACAGTGAACATCTTGGAGTG	2067
Qy	681	IleCysSerAspAlaSerAspMetGluLeuArgLeuValIleGlySerSerArgCysAla	700
Db	2068	ATCTGTGTTCTGATCGATCGGATATGAGAGCTGAGCGTGTGGTGGGAAGCAGCAGGTGCT	2127
Qy	701	GlyLysValGluValAsnValGlnGlyAlaValIleGlyLeuCysAlaAsnGlyTrpGly	720
Db	2128	GGAAAGATTGAGGTGAATGTCAGGGTGGCGTGGGAATTTCTGTGTCTAATGGCTGGGGA	2187
Qy	721	MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer	740
Db	2188	ATGAACATTCGTGAAGTTGTTTCGAGCAACTTGTAATGTGGTCTGCATCAGGGTCTCC	2247
Qy	741	ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr	760
Db	2248	AGAGAGCCTCATTTTCAGAAAGAACATACACATCTTAATGTCGAATTTCTGGCTGCAC	2307
Qy	761	GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis	780
Db	2308	GGAGGGGAAGCCTCTCTCTGGGATTGTATACGATGGAGTGGAAACACAGACTCGCTGTCA	2367
Qy	781	LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAla	800
Db	2368	TTAAATATGAAGACAGTTGATCTCTCTCAGCCACAGCCAGCCAGCTGGTGTGAGCT	2427
Qy	801	AspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys	820
Db	2428	GATATGCCCTGCTCTGGACGTGTTGAAGTGAACATGCGACACATGCGCTCTGTCTGT	2487
Qy	821	AspSerAspPheSerLeuHisIleAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp	840
Db	2488	GATTCTGATTCTCTCTTCATGCTGCCAATGTCTCTGTGCAGAGAAATTAAATTTGTGGAG	2547
Qy	841	AlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu	860
Db	2548	GCCATATCTCTTCTGTGGGAGATCACTTTGGAAAAAGGAAATGGTCTAACTTGGGCCCA	2607
Qy	861	LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro	880
Db	2608	AAGTTCCAGTGTGAAGGAGTGAACACTCACCTTGCATTTATGCCCATTTGTTCAACATCG	2667
Qy	881	GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTrpThrAspVal	900

Db 2668 GAAGACACTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTGCCGATATACAGATGTC 2727
QY 901 ArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHis 920
Db 2728 CGACTTGTGTGANTGGCAATCCAGAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACAC 2787
QY 921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940
Db 2788 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAGAAGATGCCCGTGTCTATCGACACAG 2847
QY 941 LeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerVal 960
Db 2848 CTCAGCTGTGGGAGCTGCTCTCAACACACAGGAGGAAATATATTGGAGAAAGAGTGT 2907
QY 961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGln 980
Db 2908 CGTGTGTGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTCTGGATTAAGTGTCAA 2967
QY 981 MetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThr 1000
Db 2968 ATGACAGTCTTGGAGCACTCCCTGTATCCATGGAAATACTCTCTGTGATCTGCACA 3027
QY 1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrIleuSer 1020
Db 3028 GGAGGCTCAACCCAGGCACTGTTTCCATGCCCTGCCAAATGTATCTGACCCATATTGTCT 3087
QY 1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuValAsp 1040
Db 3088 GCAGTTCCAGAGGGCAGTCTTTGATCTCTTAGAGGACAAACGGCTCCGCCCTAGTGGAT 3147
QY 1041 GlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheThrGlyThrIle 1060
Db 3148 GGGGACAGCCGCTGTCCCGGAGAGTAGAGATCTATCACAGGGCTCTCTGGGCGACCATC 3207
QY 1061 CysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080
Db 3208 TGTGATGACGGCTGGGACCTGACCGATGCCAGTGGTGTGTCAAAAGCTGGGCTGTGGA 3267
QY 1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleThrLeu 1100
Db 3268 GTGGCCTTCAATGCCACGCTCTGTCTCACTTTGGGAGGGGTTCAGGGCCCATCTGGCTG 3327
QY 1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120
Db 3328 GATGACCTGAACCTGCACAGGAAGGAGTCCCATCTTGTGCAGTGGCCCTTCCCGCGGCTGG 3387
QY 1121 GlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAla 1140
Db 3388 GGGCAGCAGCTGCAGGCACAGGAGGAGCGAGGGGTCTATCTGCTCAGAAATTCACAGCC 3447
QY 1141 LeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyr 1160
Db 3448 TTGAGGCTCTACAGTGAACCTGAAACAGAGAGTGTGCTGGGAGATTGGAAAGTCTTCTAT 3507
QY 1161 AsnGlyThrTrpGlySerValGlyArgArgAsnIleThrAlaIleAlaGlyIleVal 1180
Db 3508 AACGGACCTTGGGGACGCTGGCAGGAGGAACATCACCACAGCCATAGCAGCATTTGG 3567
QY 1181 CysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThr 1200
Db 3568 TGCAGGCAGCTGGGCTGTGGGAGATGGAGTTGTGAGCTCGCCCTTATCTAAGACA 3627
QY 1201 GlySerGlyPheMetTrpValAspAlleGlnCysProLysThrHisIleSerIleTrp 1220
Db 3628 GGCTCTGGTTTCATGTGGTGGATGATTCATTCAGTGTCTTAAACGCATATCTCCATATGG 3687
QY 1221 GlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIle 1240
Db 3688 CAGTGGCTGTCTGCCCATGGGAGCGGAAGATCTCAGGCCACAGAGAGACCTGGATC 3747
QY 1241 ThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260
Db 3748 ACATGTGAAGATAGATAGAGTGGCTGGGAGGAGACCCAGTGTCTCTGGAGAGTGGAG 3807

QY 1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280
Db 3808 ATCTGGCAGCAGGCTCTCTGGGGCACAGTGTGTGATGACTCTCTGGACCTGCCGAGGCG 3867
QY 1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300
Db 3868 GAAGTGGTGTGTGCAGCAGCTGGGCTGTGGCTCTGTCTGTCTGCTGAGGAGCCTTCG 3927
QY 1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320
Db 3928 TTTGGCCAGGGAACCTGGAAACCATCTGGTTGGATGACATCGGTGCAAGGAAATGAGTCA 3987
QY 1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340
Db 3988 TTTCTATGGGACTGTACGCCAAACCTCTGGGACAGAGTACTGTGGACACAAGGAAGAT 4047
QY 1341 AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeu 1360
Db 4048 GCTGGCGTAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCCTCTCAGGTCAATTA 4107
QY 1361 AlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPheIleLeuPheLeu 1380
Db 4108 GCACATTATTTATCCAGTATCTTTGGGCTCCTCTCTCTGTTCTGTATTCTATTCTC 4167
QY 1381 ThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArg 1400
Db 4168 ACGTGTGTCGAGATTTCAGAAACAAACATCTGCCCTCAGAGTTTCAACCAAGAGAGG 4227
QY 1401 GlySerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspPro 1420
Db 4228 GGTCTCTCGAGAGAAATTTATTCATGAGATGGAGACCTGCTCAAGAGAGAGGCCCA 4287
QY 1421 HisGlyThrArgThrSerAspAspThrProAsnHisGlyCysGluAspAlaSerAspThr 1440
Db 4288 CATGGACAAGAACCTCAGATGACACCCCAACCACTGTTGTGAAGATCTAGCGACACA 4347
QY 1441 SerLeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453
Db 4348 TCGCTGTGGGAGTTCTCTCTGCTCTGAAAGCCACAAA 4386

RESULT 3

AS01217
ID AAS01217 standard; cDNA; 3670 BP.

XX AAS01217;

XX 04-JUL-2001 (first entry)

DE DNA encoding human secreted protein, POLY8.

XX Human secreted protein; therapeutic; diagnostic; human; cancer; ss.

OS Homo sapiens.

PH Key Location/Qualifiers

FT CDS 113..3475

FT /*tag= a

FT /product= "Human secreted protein POLY8"

PN WO200119856-A2.

XX 22-MAR-2001.

XX 13-SEP-2000; 2000WO-US25106.

XX 13-SEP-1999; 99US-0153629.

PR 16-SEP-1999; 99US-0154520.

PR 20-SEP-1999; 99US-0154762.

PR 13-OCT-1999; 99US-0159231.

XX 12-SEP-2000; 2000US-0659634.

PA (CURA-) CURAGEN CORP.

XX Shimkets RA, Fernandes E, Herrmann JL, Liu X, Yang M, Boidog FL;
 XX WPI; 2001-244781/25.
 DR P-PSDB; AAU00396.
 XX
 XX New POLYX polypeptide useful for treating or preventing a POLYX
 PT associated disorder, e.g. cancer -
 PS
 PS Claim 9; Page 25-29; 152pp; English.
 XX
 XX The sequence represents the coding sequence of human secreted protein,
 CC POLY8. POLYX nucleic acids, polypeptides and antibodies to POLYX can be
 CC used for treating or preventing a POLYX associated disorder in a subject,
 CC preferably a human. These can be used in the manufacture of a medicament
 CC for treating a syndrome associated with a human disease selected from a
 CC POLYX-associated disorder, where the therapeutic is a POLYX polypeptide,
 CC a POLYX nucleotide or a POLYX antibody. They may also be used to screen
 CC for a modulator of activity, or latency, or predisposition to a POLYX
 CC associated disorder, e.g. cancer.
 XX
 XX Sequence 3670 BP; 947 A; 752 C; 1073 G; 898 T; 0 other;
 SQ

Alignment Scores:

Pred. No.:	0	Length:	3670
Score:	6107.50	Matches:	1103
Percent Similarity:	95.33%	Conservative:	0
Best Local Similarity:	95.33%	Mismatches:	3
Query Match:	75.05%	Indels:	51
DB:	22	Gaps:	1

US-09-759-130b-381 (1-1453) x AAS01217 (1-3670)

QY	348	AsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer-----	362
DB	2	AAITTTGACTGCTCTCAACAGATGCTGCTGATCTGCTCAGACCAGTTGCTATTC	61
QY	362	-----	362
DB	62	AAACTCTGGAACCTCTCGAACTATTTCCTCTGGAAGAGTAAGGAGGCCATGAAGAAT	121
QY	362	-----	362
DB	122	GATGCAGATGTGAAATAAGTTGGAAACCGTCCCAACAGTGAACAGATGTGTATGG	181
QY	363	-----AspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsn	376
DB	182	CCAACCCACAGGAAATGGATGGAGCAGATTGGAACTGGCACTAGCAGATGGAAGTAAC	241
QY	377	AsnCysSerGlyArgValGluValArgIleHisGluGlnTrpThrIleCysAspGln	396
DB	242	AAITGTTTCAGGAGAGTAGAGTGAGAATTCATGAACAGTGGTGGCAATATGTGACCAG	301
QY	397	AsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerVal	416
DB	302	AACTGGAAGAATGAACAGCCCTTGTGGTTGTAAGCAGCTAGGATGCTCGCTTCAGCGTC	361
QY	417	PheGlySerArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIle	436
DB	362	TTTGGCAGCTGCTGCTGCTAAACCTAGTAATGAAGCTAGACACATTTGGATAACAGCATA	421
QY	437	SerCysThrGlyAsnGluSerAlaLeuTrpAspCysThrTrpAspGlyLysAlaLysArg	456
DB	422	TCITGGCACTGGGAATGAGTCAGCTCTCTGGGACTGCACATATGATGAAAGCAACGGA	481
QY	457	ThrCysPheArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeu	476
DB	482	ACATGCTTCGCAAGATCAGATGCTGGAGTAATTTGTTCTGATAGGAGGATCTCGACCTA	541
QY	477	ArgLeuValGlyAlaHisSerProCysTyrClyArgLeuGluValLysTyrGlnGlyGlu	496
DB	542	AGGCTTGTGGGGCTCATAGCCCTGTTATGGGAGATTGGAGGTGAATACCAAGAGAGAG	601

QY	497	TrpGlyThrValCysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGln	516
DB	602	TGGGGGACTGTGTCTCATGACAGATGAGCAAGGAATGACGCTGTGTGTGTAACAA	661
QY	517	LeuGlyCysGlyLysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGly	536
DB	662	TTGGGATGTGGAAGCCATGATGATGTGTGTGTATGACCTATTAAAGAGACATCAGGA	721
QY	537	ProIleTrpLeuAspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGlu	556
DB	722	CCATATTTGGCTGGATGACGTTCTTTCATTCGAAATGAGTCAATATCTGGGACTGTGA	781
QY	557	HisSerGlyTrpGlyLysHisAsnCysValHisArgGluAspValIleValThrCysSer	576
DB	782	CACAGTGGATGGGAAAGCATAATTTGCTACACAGAGGATGTGATTTAACTGTCTCA	841
QY	577	GlyAspAlaThrTrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeu	596
DB	842	GGTATGCAACATGGGCGCTGAGGCTGGTGGCGGCGACACCCGCTGCTCGGGAAGACTG	901
QY	597	GluValTyrPheGlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAla	616
DB	902	GAGGTGTACTTCAAGGACGGTGGGCGACAGTGTGTGATGACGGCTGGAACGGTAAGCT	961
QY	617	AlaAlaValValCysSerGlnLeuAspCysProSerIleIleGlyMetGlyLeuGly	636
DB	962	GCAGCTGTGGTGTAGCCAGCTGGACTGCCCATCTTCTATCATTTGGCATGGGTGGGA	1021
QY	637	AsnAlaSerThrGlyTyrGlyLysIleTrpLeuAspValSerCysAspGlyAspGlu	656
DB	1022	AACGCTTCTACAGGATATGGAAATTTGGCTCGATGATGTTCTCTGTGTGGATGAG	1081
QY	657	SerAspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGlu	676
DB	1082	TCAGATCTCTGCTCATCGAGAACAGTGGTGGGGAATAATGACTGCGCATGACAGTGA	1141
QY	677	AspValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySer	696
DB	1142	CATCTTGGAGTGATCTGTTCTGATGTCATGCGATATGAGGCTGAGGCTTGTGGTGAAGC	1201
QY	697	SerArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAla	716
DB	1202	AGCAGGTGTGCTGGAAGATTGAGGTGAATGTCAGGGTGCCTGGGAATTTCTGTGTCT	1261
QY	717	AsnGlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAla	736
DB	1262	AATGGCTGGGAAGTAACATTTGCTGAACTGTTTGCAGCAACTTGAATGTGGGTCTGCA	1321
QY	737	IleArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsn	756
DB	1322	ATCAGGGTCTCCAGAGAGCCCTCAITTCACAGAAAGAACATTACACATCTTAATGTGGAAT	1381
QY	757	SerGlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGln	776
DB	1382	TCTGGCTGCATCGAGGGAGCCCTCTCTGGGATTGTATACGATGGGATGGAAACAG	1441
QY	777	ThrAlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArg	796
DB	1442	ACTCGCTGTCATTAATAATATGGAAGCAAGTTTGTATCTGCTCAGCCACAGCAGCCAG	1501
QY	797	LeuValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrp	816
DB	1502	CTGTTGGAGCTGATATGCCCTGCTCTGAGCGTGTGAAGTGAACATGCAACATGCG	1561
QY	817	ArgSerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeu	836
DB	1562	CGCTCTGTGTGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1621
QY	837	AsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeu	856
DB	1622	ACTGTGGAGATGCATATCTTCTTCTGAGAGATCATCTTGGAAAGGAAGTGGTCTA	1681
QY	857	ThrTrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIle	876

Db 1682 ACTTGGCCGAAAGTTCAGTGTGAAGGAGTGAACATCACTTGCATTATGCCCCATT 1741
Qy 877 ValGlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArg 896
Db 1742 GTTCAACATCCGAGACACTGTATCCACACAGAGAAGTTGGAGTGTCTGTGCCGA 1801
Qy 897 TyrThrAspValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluLeuAsn 916
Db 1802 TATACAGATGTCGAGTCTGTGAATGGCAATCCAGTGTGACGGGCAAGTGGAGATCAAC 1861
Qy 917 ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal 936
Db 1862 GTGCTGGACATGGGGCTCACTGTGTGACACCCACTGGGAGCCAGAAAGATCCCGTGT 1921
Qy 937 LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrGlyGlyLysTrpIleGly 956
Db 1922 CTATGCAGACAGCTCAGCTGTGGACTGCTCTCAACACAGAGGAGAAATATATGGA 1981
Qy 957 GluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu 976
Db 1982 GAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCG 2041
Qy 977 AspAsnCysGlnMetThrValLeuGlyAlaProCysIleHisGlyAsnThrValSer 996
Db 2042 GATACTGTCAATGACAGTCTTGAGGACCTCCCTGTATCCATGGAAATACTGTCTCT 2101
Qy 997 ValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAsp 1016
Db 2102 GTGATCTGCACAGGAAGCTGACCAGCCACTGTTTCCATGCTCGCAATGTATCTGAC 2161
Qy 1017 ProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeu 1036
Db 2162 CCATATTTCTCTGCAGTCCAGAGGCGAGTGTTCATCTGCTGTAGAGGACAAAGCGCTC 2221
Qy 1037 ArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleThrHisAspGlyPhe 1056
Db 2222 CGCCTAGTGGATGGGACAGCGCTGTGCCGGAGATGAGATCTATCAGCAGCGCTC 2281
Qy 1057 TrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLys 1076
Db 2282 TGGGGCACCATCTGTGATCAGCGCTGGGACCTGAGCGATGCCAGTGGTGTGTCAAAAG 2341
Qy 1077 LeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGly 1096
Db 2342 CTGGGCTGTGAGTGGCCCTCAATGCCAGGCTCTGCTCACTTTGGGGAGGGGTGAGG 2401
Qy 1097 ProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysPro 1116
Db 2402 CCCATCTGGCTGGATGACCTGAACCTGCACAGGAATGGAGTCCCATCTGTGGCAGTGCCCT 2461
Qy 1117 SerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSer 1136
Db 2462 TCCCGCGCTGGGGCAGCAGCAGCTGCAGCCACAGGAGGAGCGAGGGTCACTCTCTCA 2521
Qy 1137 GluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeu 1156
Db 2522 GAATTCACAGCCTTGAAGTCTTACAGTGAACACTGAAACAGGAGGCTGTCTGGGAGATTG 2581
Qy 1157 GluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIle 1176
Db 2582 GAAGTCTTCTATAAGCGGACCTGGGGCAGCGTGGCAGGAGGAACATCACACAGCCATPA 2641
Qy 1177 AlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaPro 1196
Db 2642 GCAGCATTTGTGCAGGAGCTGGGCTGTGGGGAGATGGAGTTGTACGCTCGCCCT 2701
Qy 1197 LeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHis 1216
Db 2702 TTATCTAAGACAGCGCTCTGGTTTCTATGTGGTGGATGACATTCAGTGTCTTAAACGCAT 2761
Qy 1217 IleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGlu 1236

Db 2762 ATCTCCATATGCGAGTGCCTGTCTGCCCCCATGGGAGCGAAGAATCTCCAGCCGAGAGAA 2821
Qy 1237 GluThrTrpIleThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSer 1256
Db 2822 GAGACTGGATCACATGTCAAGATAGAATAAGAGTGCCTGGAGGAGACACCGAGTCTCT 2881
Qy 1257 GlyArgValGluIleThrPheHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAsp 1276
Db 2882 GGGAGAGTGGAGATCTGGCAGCAGGCTCTCTGGGGCACAGTGTGTGATGACTCTCTGGGAC 2941
Qy 1277 LeuAlaGluAlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeu 1296
Db 2942 CTGGCCGAGCGGAAGTGTGTGTGACAGCTGGGCTGTGCTCTCTGGTGGCCCTG 3001
Qy 1297 ArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrpLeuAspMetArgCysLys 1316
Db 3002 AGGGAGCGTTCGTTTGGCCAGGGAACCTGGAACCATCTGTTGGATGACATGCGGTGCAAA 3061
Qy 1317 GlyAsnGluSerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGly 1336
Db 3062 GGAAATGAGTCAATTTCTATGGGAGTGTACAGCCAAACCTGGGGACAGAGTGA 3121
Qy 1337 HisLysGluAspAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSer 1356
Db 3122 CACAAGGAAGATGCTGGCGTGAGGTCTCTGGACAGTCTGAAATCACTGAATGCTCC 3181
Qy 1357 SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPhe 1376
Db 3182 TCAGGTCAATTTAGCACATTTATTTATCCAGTATCTTTGGGCTCTCTCTCTGTTCTGTTT 3241
Qy 1377 IleLeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSer 1396
Db 3242 ATTCTATTTCTACGTGGTGGCGAGTTCAGAAACAAAAACATCTGCCCTCTCAGAGTTCA 3301
Qy 1397 ThrArgArgGlySerLeuGluAsnLeuPheHisGluMetGluThrCysLeuLys 1416
Db 3302 ACCAGAGAGGGGTCTCTCGAGGAGAAATTTATTCATGAGATGGAGACCTGCCCTCAAG 3361
Qy 1417 ArgGluAspProHisGlyThrArgThrSerAspAspThrProAsnHisGlyCysGluAsp 1436
Db 3362 AGAGAGGACCCACATGGGACAAAGAACCTCAGATGACACCCCAACCATGTTGTGAAGAT 3421
Qy 1437 AlaSerAspThrSerLeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453
Db 3422 GTAGCGACACATCGCTGTTGGAGTCTTCTCTGCTCTGAAAGCCACAAA 3472
RESULT 4
AAS44609
ID AAS44609 standard; DNA; 2800 BP.
XX
AC AAS44609;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human full-length polynucleotide sequence #34.
XX
KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
KW nervous system disorder; inflammatory disorder; cell differentiation; ds;
KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
KW cytostatic; antirheumatic; antilithritic; vulnary; antinflammatory;
KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
KW immunostimulant; analgesic; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200164834-A2.
XX
PD 07-SEP-2001.
XX

PF 26-FEB-2001: 2001WO-US04926.
 XX 28-FEB-2000: 2000US-0515126.
 PR 18-MAY-2000: 2000US-0577409.
 PR 17-JUN-2000: 2000US-0597707.
 PR 14-JUL-2000: 2000US-0616807.
 PR 19-SEP-2000: 2000US-0664641.
 XX (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
 PI Drmanac R;
 XX
 XX WPI: 2001-589862/66.
 DR P-PSDB; AAU27709.
 XX
 PT Novel polypeptides and nucleic acids obtained from cDNA libraries
 prepared from various human tissues, for diagnosis, treatment of
 PT cancer, neurological, inflammatory disorders and for use in arrays for
 PT detection
 XX
 PS Claim 1: SEQ ID NO 34; 153pp; English.
 XX
 CC Sequences AAS44576-AAS44919 represent full-length polynucleotides and
 CC contig polynucleotides encoding polypeptides of the invention. The DNA
 CC and protein sequences are useful for the treatment, diagnosis and
 CC prevention of various types of disorder in a mammalian subject such as a
 CC human, dog, monkey, mouse, hamster or rat. The disorders include cancers
 CC such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
 CC as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
 CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
 CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
 CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
 CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
 CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
 CC bowel disease. The sequences exhibit activity relating to angiogenesis,
 CC cell proliferation, cell differentiation, stem cell growth factor,
 CC activin or inhibin. Therefore, they can be used to manipulate stem cells
 CC in culture to give rise to neuroepithelial cells that can be used to
 CC augment or replace cells damaged by illness, accidental damage or genetic
 CC disorders. The sequences may also be used for regeneration of bone,
 CC cartilage, tendons and ligaments and in tissue repair and burn healing.
 CC Note: Some sequences for this patent did not form part of the printed
 CC specification, but were obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 2800 BP; 697 A; 607 C; 813 G; 683 T; 0 other;

Alignment Scores:
 Pred. No.: 3.57e-293 Length: 2800
 Score: 4712.00 Matches: 850
 Percent Similarity: 99.77% Conservativity: 0
 Best Local Similarity: 99.77% Mismatches: 2
 Query Match: 57.90% Indels: 0
 DB: 22 Gaps: 0

US-09-759-130B-381 (1-1453) x AAS44609 (1-2800)

Qy 602 GlyArgTrpGlyThrValCysAspGlyTrpAsnSerLysAlaAlaValValCys 621
 Db 37 GCACGGTGGGCGACAGCGGTGATGACGGCTGGAAACAGTAAGCTGCAGCTGGTGTGT 96
 Qy 622 SerGlnLeuAspCysProSerSerIleGlyMetGlyLeuGlyAsnAlaSerThrGly 641
 Db 97 AGCCAGCTGGAGTCCCATCTCTATCATTTGGCATGGCTCTGGGAAAGCTCTACAGGA 156
 Qy 642 TyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSer 661
 Db 157 TATGAAAAATTTGGCTCGATGATGTTCTCTGTCATGGAGATGATGATGATCTCTGGTCA 216
 Qy 662 CysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValIleValIle 681

Db 217 TGCAGGAACACTGGGTGGGAAATAATCACTGCAGTCACAGTGAAGATGTTGGAGTGATC 276
 Qy 682 CysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGly 701
 Db 277 TGTTCATGATCGATATGAGCTGAGGCTTGTGGGTGGGAAGCAGCAGGTGCTCGA 336
 Qy 702 LysValGluValAsnValGlnGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMet 721
 Db 337 AAAGTTGAGGTGAATGTCAGGGTGCCTGGGAATTCCTGTCTAATGGCTGGGAATG 396
 Qy 722 AsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArg 741
 Db 397 AACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGTCTGCAATCAGGCTCTCCAGA 456
 Qy 742 GluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGly 761
 Db 457 GAGCCTCATTTACAGAAAGAACATTTACATCTTAATGTCTGAATTCCTGGCTGCACTGA 516
 Qy 762 GlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeu 781
 Db 517 GGGGAAGCCTCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGCATTTA 576
 Qy 782 AsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAsp 801
 Db 577 AATATGGAAGCAAGTTTGTATCTGCTCAGCCACAGCCAGCCAGGCTGGTTGGAGCTGAT 636
 Qy 802 MetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAsp 821
 Db 637 ATGCCCTGCTCTGACGCTGTTGAAGTGAACATGCAGACACATGCGGCTCTGCTGTGAT 696
 Qy 822 SerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAla 841
 Db 697 TCTGATTTCTCTCTCATGCTGCCAATGCTGTGCAGAGAATTAACCTGTGGAGATGCC 756
 Qy 842 IleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLys 861
 Db 757 ATATCTCTTTCTGGGGAGATCACTTTGGAAAGGAATGTTCTAACTTGGGCCGAAAG 816
 Qy 862 PheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGlu 881
 Db 817 TTCAGTGTGAAGGAGTGAACCTCACCTTGCATTTATCCCATTTGTTCAACATCCGAA 876
 Qy 882 AspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspValArg 901
 Db 877 GACACTTGTATCCACAGCAGAGAAGTTGGAGTTGCTCTCCGATATACAGATGTCGA 936
 Qy 902 LeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrp 921
 Db 937 CTTGTGAATGGCAATCCAGTGTGACGGCAAGTGGAGATCAACGTGCTGGACACTGG 996
 Qy 922 GlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeu 941
 Db 997 GGCTCACTGTGTGACACCCACTGGGACCCAGAGAATGCCGTGTTCTATGCAGACAGCTC 1056
 Qy 942 SerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArg 961
 Db 1057 AGCTGTGGGACTGCTCTCTCAACACAGCAGGAGAAATATATTGGAGAAAGAGTGTTCGT 1116
 Qy 962 ValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMet 981
 Db 1117 GTGGGGACACAGAGTTTCATTTGTTAGGGAATAGTCACTTCTGGATAACTGTCAAAATG 1176
 Qy 982 ThrValLeuGlyAlaProCysIleHisGlyAsnThrValSerValIleCysThrGly 1001
 Db 1177 ACAGTCTTGGAGCACTCCCTGTATCCATGGAATACTGCTCTGTGATGTCACAGGA 1236
 Qy 1002 SerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAla 1021
 Db 1237 AGCCTGACCCAGCCACTGTTTCCATGCTCGCAAAATGATCTGACCCCATATTGTTCTGCA 1296
 Qy 1022 ValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGly 1041
 Db 1297 GTTCCAGAGGCGTGTCTTGTATCTGCTTAGAGGACAAACGGCTCCGCTAGTGTGATGGG 1356

QY 1042 AspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCys 1061
DB 1357 GACACCGCTGTGCGGAGAGTAGAGATCATACAGCGGCTTCTGGGACCATCTGT 1416
QY 1062 AspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyVal 1081
DB 1417 GAIGACGGCTGGACCTGAGCGATGCCACGTTGGTGTCAAAAGCTGGGCTGTGGAGTG 1476
QY 1082 AlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAsp 1101
DB 1477 GCCTTCAATGCCACGGTCTCTGCTACATTGGGAGGGGTGAGGCCCATCTGGCTGGAT 1536
QY 1102 AspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGly 1121
DB 1537 GACCTGAATGCACAGGAATGAGTCCACATTGTGGCAGTGCCTTCCCGGCTGGGG 1596
QY 1122 GlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeu 1141
DB 1597 CAGCAGGACTGCAGGCACAAGGAGGACGAGGGGTCTCTGCTCAGAATTTCAGCCCTG 1656
QY 1142 ArgLeuTrpSerGluThrGluSerCysAlaGlyArgLeuGluValPheTrpAsn 1161
DB 1657 AGGCTCTACAGTGAACCTGAACACAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAC 1716
QY 1162 GlyThrTrpGlySerValGlyArgArgAsnIleThrAlaIleAlaGlyIleValCys 1181
DB 1717 GGGACCTGGGCGAGCTGGCAGGAGGAGAACATCACACAGCCATGACGAGCATTTGTGC 1776
QY 1182 ArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGly 1201
DB 1777 AGGAGCTGGGCTGTGGGAGGAATGGAGTTGTACGCTCGCCCTTTATCTAAGACAGC 1836
QY 1202 SerGlyPheMetTrpValAspIleGlnCysProLysThrHisIleSerIleTrpGln 1221
DB 1837 TCTGGTTTCATGTGGGTGAGTACATTGAGTCTCTAAACCATATCTCCATATGGCAG 1896
QY 1222 CysLeuSerAlaProTrpGluArgIleSerSerProAlaGluGluThrTrpIleThr 1241
DB 1897 TGCTGTCTGCCCTTGGGAGGAAGATCTCCAGCCAGCAGAGACCTGGATCACA 1956
QY 1242 CysGluAspArgIleArgValArgGlyArgGlyAspThrGluCysSerGlyArgValGluIle 1261
DB 1957 TGTGAAGATAGAAATAGAGTGGCTGGAGAGACACCGAGTGTCTGGGAGAGTGGAGATC 2016
QY 1262 TrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGlu 1281
DB 2017 TGGCAGCAGGCTCTGGGACACAGTGTGTGATGACTCTGGGACCTGGCCGAGCGGAA 2076
QY 1282 ValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPhe 1301
DB 2077 GTGGTGTGTACAGCAGTGGCTGTGGCTGTCTGTGGCTGTGGGACGCTTCGTTT 2136
QY 1302 GlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPhe 1321
DB 2137 GCGCAGGGAACCTGGAAACCATCTGGTGGATGACATGCGGTGCAAGGAAATGAGTCATT 2196
QY 1322 LeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAla 1341
DB 2197 CTATGGGACTGTACGCCAACCCCTGGGACAGAGTGTGAGACACACAGGAAGATGCT 2256
QY 1342 GlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeuAla 1361
DB 2257 GCGGTGAGGTGCTGTGGCAGTGTGTAATCACTGAATGCTCTGCTGAGTCAATTTAGCA 2316
QY 1362 LeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPheIleLeuPheLeuThr 1381
DB 2317 CTATTTTATCCAGATCTTTGGGCTCCTTCTCCCTGGTGTCTGTTTATTTCTATTTCTCAG 2376
QY 1382 TrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerTrpArgArgGly 1401
DB 2377 TGGTCCGAGTTCAGAAACAAAACATCTGCCCTCAGAGTTTCAACAGAGAGGGGT 2436

QY 1402 SerLeuGluGluAspLeuPheHisGluMetGluThrCysLeuLysArgGluAspProHis 1421
DB 2437 TCTCTCGAGGAGAAATTTATTCATGAGATGGAGAGCTGCTCAAGAGAGAGACCCACAT 2496
QY 1422 GlyThrArgThrSerAspAspThrProAsnHisGlyCysGluAspAlaSerAspThrSer 1441
DB 2497 GGGACAAAGACCTCAGATGACACCCCAACCATGTTGTGAAGATGCTAGCGACACATCG 2556
QY 1442 LeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453
DB 2557 CTGTGGGAGTCTTCTCTGCTCTGAAAGCCACAAAA 2592
RESULT 5
AAI58649
ID AAI58649 standard; cDNA; 3716 BP.
XX
AC AAI58649;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 852.
XX
KW Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM39493.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
PS Claim 1; SEQ ID NO 852; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with neotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localized neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and

CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed
CC specification.XX
SQ Sequence 3716 BP; 993 A; 730 C; 1060 G; 933 T; 0 other;

Alignment Scores:

Pred. No.:	1, 2e-210	Length:	3716
Score:	3429.00	Matches:	598
Percent Similarity:	71.27%	Conservative:	161
Best Local Similarity:	56.15%	Mismatches:	294
Query Match:	42.14%	Indels:	12
DB:	22	Gaps:	4

US-09-759-130B-381 (1-1453) x AA158649 (1-3716)

Qy 364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383
 Db 222 GGAACAGACAAGAGCTGAGCTAGTGGATGGTGAACAACAAAGTGTAGCGGAGAGTGGAA 281
 Qy 384 ValArgIleHisGluGlnTrpThrIleCysAspGlnAsnTrpLysAsnGluAla 403
 Db 282 GTGAAGTCCAGGAGGAGTGGGACCGGTGTATTAATGCTGGACCATGGAACGGCTC 341
 Qy 404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgAlaLys 423
 Db 342 TCTGTGATTTGTAACAGCTGGGATGCCAACTGCTATCAAGCCCTGGATGGCTAAT 401
 Qy 424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443
 Db 402 TCCAGTGCAGGTTCTGGACGCAATTGGATGGATCATGTTCTTCTGCGTGGGAATGAGTCA 461
 Qy 444 AlaLeuTrpAspCysThrTyTrpAspGlyLysAlaLysArgThr--CysPheArgArgSer 462
 Db 462 GCTCTTGGGATTCGAACATGATGGATGGGAAAGCATAGTAAGTACTGACTCACCACAA 521
 Qy 463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis 482
 Db 522 GATGCTGGAGTGACCTGCTAGATGGATGCCAATTTGGAAATGAGGCTGACCGTGGAGGG 581
 Qy 483 SerProCysTyTrpGlyArgLeuGluValLysTyTrpGlnGlyGluTrpGlyValCysHis 502
 Db 582 AATATGTCTTGGAAAGTAATAGATCAATTCGAAGCGGTGGGAAACAGTGTGTGAT 641
 Qy 503 AspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPro 522
 Db 642 GATAACTCAACATAGATCATGCTGCTCATTTGTAGACAACTGAATGTGGAAGTCT 701
 Qy 523 MetHisValPheGlyMetThrTyTrpPheLysGluAlaSerGlyProIleTrpLeuAspAsp 542
 Db 702 GTCAGTTTCTCTGTTTCATCTAAATTTGGAGAAGGCTCTGGACCAATCTGTTTATGATGAT 761
 Qy 543 ValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys 562
 Db 762 CTTATATGCAACGGAATGATGACGCTCTGGAAGTCAACATCAAGAGTGGGGAAG 821
 Qy 563 HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly 582
 Db 822 CATNACTGTGATCATGCTGAGGATGCTCGAGTATTGCTCAAGGGAGCATCTGAGC 881
 Qy 583 LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyTrpPheGlnGly 602
 Db 882 CTGAGACTGGTAGATGATGATGACGGCTGGACAGATTACGATGCTGTGGCATGCAAG 1001
 Qy 603 ArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValValCysSer 622
 Db 942 GAATGGGGGCAATATGATGACGGCTGGACAGATTACGATGCTGTGGCATGCAAG 1001
 Qy 623 GlnLeuAspCysProSerIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyTrp 642
 Db 1002 CAACTGGGATGTCACNCTGCCGTGCACAGCCATGTGCTGCAAGTAAAGGATTT 1061
 Qy 643 GlyLysIleTrpLeuAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys 662

Db 1062 GGACACATCTGGCTTGACAGCGTTTCTGCCAGGACATGAACCTGCTGTGGCAATGT 1121
 Qy 663 ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys 682
 Db 1122 AACACCATGAATGGGAAAGCATATTATTCGAATCAATGAAGATGCTGGCGTGACATGT 1181
 Qy 683 SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLys 702
 Db 1182 TCTGATGATCAGATCTGGACCTAGAGCTTAGAGGTGGAGCAGCGCTGTGCTGGGACA 1241
 Qy 703 ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn 722
 Db 1242 GTTGAGGTGGAGATTACAGACTGTTAGGGAAGGTGTGTGACAGAGCGTGGGACTGAAA 1301
 Qy 723 IleAlaGluValValCysArgGlnLeuValCysGlySerAlaIleArgValSerArgGlu 742
 Db 1302 GAAGCTGATGTTGGTTGAGGACGCTGGGATGGATGTCACCTCAAAACATCTTATCAA 1361
 Qy 743 ProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGly 762
 Db 1362 GTGTACTCCAAATCCAGGCAACAAACACATGGCTGTTCTTAAGTAGCTGTAAACGAAAT 1421
 Qy 763 GluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsn 782
 Db 1422 GAAACTTCTCTTTGGGACTGCAAGAACTGGCAATGGGTGGACTTACCTGTGATCCTAT 1481
 Qy 783 MetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMet 802
 Db 1482 GAAGAAGCAAAATTACCTGCTCAGCCCAAGGAAACCCAGACTGGTTGGAGGGACAAAT 1541
 Qy 803 ProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSer 822
 Db 1542 CCCTGTTCTGACGCTGTTGAAGTGAAGCATGTTGACACGTGGGCTCCATCTGTGATTCG 1601
 Qy 823 AspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIle 842
 Db 1602 CACTTCTCTCTGGAAGCTGCCAGCGTTCTATGCAAGGAAATTACAGTGTGGCACAGTTC 1661
 Qy 843 SerLeuSerValGlyAspPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPhe 862
 Db 1662 TCTATCTCTGGGGAGGACTCCTCTTGGAGAGGAAATGACAGACTGGGCTGGCAAGATTC 1721
 Qy 863 GlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp 882
 Db 1722 CAGTGTGAGGACATGATGCCATCTTCTCACTCTGCCAGTAGCACCCGCCCAAGAGGA 1781
 Qy 883 ThrCysIleHisSerArgGluValGlyValValCysSerArgTyTrpThrAspValArgLeu 902
 Db 1782 ACTTGTAGCCACAGCAGGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTG 1841
 Qy 903 ValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGly 922
 Db 1842 GTGAATGCAAGACCCCTGTGAGGGCAGTGGAGCTCAAAACCTTGGTCCCTGGGGA 1901
 Qy 923 SerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSer 942
 Db 1902 TCCCTCTCTAACTCTCACTGGGACATAGAAGATGCCATGTTCTTTGGCAGCAGCTTAA 1961
 Qy 943 CysGlyThrAlaLeuSerThrThrGlyGlyLysTyTrpIleGlyGluArgSerValArgVal 962
 Db 1962 TGTGAGTGTGCCCTTCTTACCCAGGAGGACACCGTTTGGAAAAAGGAAATGTCAGATC 2021
 Qy 963 TrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThr 982
 Db 2022 TGGAGGCATATGTTTCACTGCACTGGGACATGAGCAGCAGATGGGAGATGTCCTTA 2081
 Qy 983 ValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGlySer 1002
 Db 2082 GCTCTAGTGTCTTATGATGCTTCCAGCAAGTGGGCTCTGTATNTCTGCTCAGGAAC 2141
 Qy 1003 LeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyTrpLeuSerAlaVal 1022

Db 2142 CAGTCCCAACACTGCTCTGTCGCAATTCATCGTCTTTGGGCCCAACAGGCGCTACCAATT 2201
Qy 1023 ProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAsp 1042
Db 2202 CCAGAAAGAAGTGTGGCCCTGCATAGAGAGTGTCTCACTTGGCTGTGAATGGAGGA 2261
Qy 1043 SerArgCysAlaGlyArgValGluIleThrHisAspGlyPheTrpGlyThrIleCysAsp 1062
Db 2262 GGTGCGTGTGCTGGGAGAGTATCATATGAGGGCTCTGGGCGACCATCTGTGAT 2321
Qy 1063 AspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAla 1082
Db 2322 GACAGCTGGGACCTGAGTATGCCACGTGTTGTCAGACAGCTGGGCTGTGGAGAGGCC 2381
Qy 1083 PheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspAsp 1102
Db 2382 ATTAATGCCACGTGTTCTGCTCATTTTGGGAGAGGACAGGCGCCATCTGGCTGGATGAG 2441
Qy 1103 LeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGln 1122
Db 2442 ATGAATGCAATGGAAGAATCCCGCAATTGGCAGTGCCATTCACACGGCTGGGGCGAG 2501
Qy 1123 HisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArg 1142
Db 2502 CAAATTTGAGGCACAGAGAGTATGCTGCTCAGAAATCATGTCTCTGAGA 2561
Qy 1143 LeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGly 1162
Db 2562 CTACACAGTGAACCCAGCAGAGAGGCTGTGCGAGGCGCTGTGGAAGTTTTTACATGGA 2621
Qy 1163 ThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArg 1182
Db 2622 GCITGGGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGTGTGGTGTGCAGG 2681
Qy 1183 GlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySer 1202
Db 2682 CAGCTGGGCTGTGCACAAAGGAAATCAACCCCTGCATCTTTACAGAGGCCATGTCTCC 2741
Qy 1203 GlyPheMetTrpValAspAlaGlnCysProLysThrHisIleSerIleTrpGlnCys 1222
Db 2742 ATTCCTCATGTGGTGACAAATGTTAGTGTCCAAAGGACCTGACACGCTGTGGCAGTGC 2801
Qy 1223 LeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleThrCys 1242
Db 2802 CCATCATCTCCATGGGAGAGAGACTGGCCAGCCCTCGGAGGAGACCTGGATCATCATGT 2861
Qy 1243 GluAspArgIleArgValArgGlyAspThrGluCysSerGlyArgValGluIleTrp 1262
Db 2862 GACAACAAGATAAGACTTCAGGAGAGACCCACTTCTGTTGAGACGTGTGGAGATCTGG 2921
Qy 1263 HisAlaGlySerTrpGlyThrValCysAspSerTrpAspLeuAlaGluAlaGluVal 1282
Db 2922 CATGGAGTCTCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTG 2981
Qy 1283 ValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGly 1302
Db 2982 GTGTGTCAACAACCTGTGCTGTGCTCAGCTTTGAAAGCATTTCAAAAGCAGAGAGTGTGGT 3041
Qy 1303 GlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeu 1322
Db 3042 CAGGGGACTGGCCGATATGGCTCAATGAAGTGAAGTGAAGGAAATGATGATCTTCTTC 3101
Qy 1323 TrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGly 1342
Db 3102 TGGGATGTCTGCCAGACGCTGGGGCCATAGTGTGAGTGTGGGCACAAAGGAGACGCTGCA 3161
Qy 1343 ValArgCysSerGlyGlnSerLeuLys-----SerLeuAsnAlaSer 1356
Db 3162 GTGAATTTGCACAGATATTTTCAGTGCAGAAACCCCAAAAGCCACACAGCGTGCCTCA 3221
Qy 1357 SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPhe 1376
Db 3222 TCCCGTCAGTCATCTTATTTGTCAGTCGGGATCTCTTGGGTTGTTGTTGGCCATTTTC 3281

Qy 1377 Ile-----LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArg 1394
Db 3282 GTCGCATTATCTTCTTCTGACTAAAAAGCGAAGACAGACAGACGCGG-----CTTGCA 3332
Qy 1395 ValSerThrArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCys 1414
Db 3333 GTTTCCTCAAGAGGAGAGAACTTAGTCCACCAATTCATACCGGGAGATGAATTTCTTGC 3392
Qy 1415 LeuLysArgGluAsp 1419
Db 3393 CTGAATGCAGATGAT 3407
RESULT 6
ID AAH98740 standard; cDNA; 3834 BP.
XX AAH98740;
AC AAH98740;
XX 12-OCT-2001 (first entry)
XX Human EST-derived coding sequence SEQ ID NO: 597.
DE Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder;
KW biodiversity; gene therapy; nutrition; ss.
OS Homo sapiens.
XX WO200154477-A2.
XX 02-AUG-2001.
XX 25-JAN-2001; 2001WO-US02687.
XX 25-JAN-2000; 2000US-0491404.
PR 17-JUL-2000; 2000US-0617746.
PR 03-AUG-2000; 2000US-0631451.
PR 15-SEP-2000; 2000US-0663870.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Drmanac RA, Zhang J, Wehrman T;
XX WPI; 2001-476164/51.
DR P-PSDB; AAM24081.
XX Isolated polypeptide for treatment of diseases, diagnostics, raising
PT antibodies and research use -
PS Claim 1; Page 584-585; 1275pp; English.
XX The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a cDNA
CC of the invention.
XX Sequence 3834 BP; 1017 A; 756 C; 1082 G; 979 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 1,25e-210 Length: 3834
Score: 3429.00 Matches: 598
Percent Similarity: 71.27% Conservative: 161
Best Local Similarity: 56.15% Mismatches: 294
Query Match: 42.14% Indels: 12
DB: 22 Gaps: 4

QY 1083 PheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleThrLeuAspasp 1102
DB 2382 ATTAATGCCACCTGGTTCTCTCTATTTGGGGAAGAACAGGCCCATCTGGCTGGATGAG 2441
QY 1103 LeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGln 1122
DB 2442 ATGAATGCAATGGAAAGAAATCCCGCATTTGGCAGTGCATTCACAGCGCTGGGGCAG 2501
QY 1123 HisAsnCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArg 1142
DB 2502 CAAATTTGAGGACACAGGAGATGGGGAGTATCTGCTCAGAAATTCATGCTCTGTGAGA 2561
QY 1143 LeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGly 1162
DB 2562 CTGACAGTGAAGCCACGACAGAGAGCGCTGCGAGGCGCTCTGGAAGTTTTTACATGGA 2621
QY 1163 ThrTrpGlySerValGlyArgArgAsnIleThrAlaIleAlaGlyIleValCysArg 1182
DB 2622 GCTTGGGCGACTGTGGCAAGAGTAGCATGCTGAAACCACCTGTGGGTGTGGTGGCAGG 2681
QY 1183 GlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySer 1202
DB 2682 CAGCTGGCTGTGCAGACAAAGGAAATCAACCTGTCATCTTAGACAAAGCCCATGTCC 2741
QY 1203 GlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCys 1222
DB 2742 ATTCATCTGGTGGTGGACATGTTTCAGTGTCCAAAGGACCTGCACAGCTGTGGCAGTCC 2801
QY 1223 LeuSerAlaProTrpGluArgGlySerProAlaGluThrTrpIleThrCys 1242
DB 2802 CCATCATCTCCATGGAGAGAGAGCTGGCCAGCCCTCGGAGAGACCTGGATCATGTC 2861
QY 1243 GluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrp 1262
DB 2862 GACACAAAGATAAGACTTCAGGAAGAGACCCACTTCTGTTCTGGAGCTGTGGAGATCTCG 2921
QY 1263 HisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaVal 1282
DB 2922 CATGGAGTTCCTGGGGGACAGTGTGTAGTACTCTGGGACTTGGACGATGCTCAGGTG 2981
QY 1283 ValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGly 1302
DB 2982 GTGTGTCAACAACTGGCTGTGTCCAGCTTTGAAAGCATTCAAAGAACAGCAGATTGCT 3041
QY 1303 GlnGlyThrGlyThrIleTrpLeuAspMetArgCysLysGlyAsnGluSerPheLeu 1322
DB 3042 CAGGGGACTGGACCGATATGGCTCAATGAAGTGAAGTCAAAAGGAATGAGTCTTCTCTG 3101
QY 1323 TrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGly 1342
DB 3102 TGGGATGTCTGCGCAGAGCTGGGGCCATAGTGTGTGGCAGCAAGGAACAGCCTGCA 3161
QY 1343 ValArgCysSerGlyGlnSerLeuLys-----SerLeuAsnAlaSer 1356
DB 3162 GTGAATTCACAGATATTCAGTGCAGAAACCCCAAAAGCCCAACAGCGTCCCTCA 3221
QY 1357 SerGlyHisLeuAlaLeuIleLeuSerIlePheGlyLeuLeuLeuValLeuPhe 1376
DB 3222 TCCCGTCAGTCATCTTTATTTCAGTCGGGATCCCTTGGGGTGTCTGTGTGGCCATTTTC 3281
QY 1377 Ile-----LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArg 1394
DB 3282 GTCCGATTTCTCTGTACTAAAGCGAAGACAGACAGACGCG-----CTTGCA 3332
QY 1395 ValSerThrArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCys 1414
DB 3333 GTTTCCTCAAGGAGAGAACTTAGTCCACCAAAATTCATACCGGGAGATGAATCTTCTTC 3392
QY 1415 LeuLysArgGluasp 1419
DB 3393 CTGAATGCAGATGAT 3407
RESULT 7

AAI60435/c
ID AAI60435 standard; cDNA; 3834 BP.
XX
AC AAI60435;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4424.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Dmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM41279.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 4424; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 3834 BP; 979 A; 1082 C; 756 G; 1017 T; 0 other;

Alignment Scores:
Pred. No.: 1,25e-210 Length: 3834
Score: 3429.00 Matches: 598
Percent Similarity: 71.27% Conservative: 161
Best Local Similarity: 56.15% Mismatches: 294
Query Match: 42.14% Indels: 12
DB: 22 Gaps: 4

US-09-759-1308-381 (1-1453) x AAI60435 (1-3834)

Qy	364	GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu	383
Db	3613	GGACACAGAGGAGCTGAGCTAGTGGTGAACAACAGTGTAGCGGAGAGTGAA	3554
Qy	384	ValArgIleHisGluInTrpThrIleCysAspGlnAsnTrpLysAsnGluAla	403
Db	3553	GTGAAGTCCAGGAGAGTGGGACCGTGTATATATGCTGCGAGCATGGAACGGCTC	3494
Qy	404	LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys	423
Db	3493	TCTGTGATTTGTACACAGCTGGGATGCCAAGTCTATCAAGGCCCTGGATGGGCTAAT	3434
Qy	424	ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer	443
Db	3433	TCCAGTCAGGTTCTGGACGATTTGGATGATCATGTTCTTGTGCGGGAATGAGTCA	3374
Qy	444	AlaLeuTrpAspCysThrTyArgPheLysAlaLysArgThr---CysPheArgArgSer	462
Db	3373	GCTCTTTGGGATTGCCAACATGATGGATGGGGAAGCATATACTGTACTCACCAACA	3314
Qy	463	AspAlaGlyValIleCysSerAspLysAlaAspLeuArgLeuValGlyAlaHis	482
Db	3313	GATCGTGGAGTACGCTGCTAGATGGATGCCAATTTGGAAATGAGCTGACCGCTGGAGG	3254
Qy	483	SerProCysTyArgLeuGluValLysTyArgGlnGlyGluTrpGlyThrValCysHis	502
Db	3253	AATATGTCTTGGGAAGATAGATCAATTCGAAGGACGGTGGGGAACAGTGTGTAT	3194
Qy	503	AspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPro	522
Db	3193	GATACTTCACATAGATCATGCTGCTATTTGTAGACAACTTGAATGTGGAAGTCT	3134
Qy	523	MethHisValPheGlyMetThrTyPheLysGluAlaSerGlyProIleTrpLeuAspAsp	542
Db	3133	GTCAGTTCTCTGTTTCATCTAATTTTGGGAAGGCTCTGGACCAATCTGTTGATGAT	3074
Qy	543	ValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys	562
Db	3073	CTTATATGCAACGGAATGATGATGCTGGAACATCAACAGTCAAGGATGGGGAAG	3014
Qy	563	HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly	582
Db	3013	CATAACTGTGATCATGCTGAGGATGCTGGAGTATTTGCTCAAGGGAGCAGATCTGAGC	2954
Qy	583	LeuArgLeuValGlyLysSerAsnArgCysSerGlyArgLeuGluValTyPheGlnGly	602
Db	2953	CTGAGATGCTGATGAGTGGATGCTCACTCAATTTTGGGAAGATTTAGAAGTTCACAGGA	2894
Qy	603	ArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValValCysSer	622
Db	2893	GAATGGGGGCAATATGTTGATGACGGCTGGGACAGTACAGATGCTGCTGGCATGCAAG	2834
Qy	623	GlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTy	642
Db	2833	CAACTGGGATTCCAACTGCGCTCACACCATTTGCTCAGTTTAACCCAGTAAGGATTT	2774
Qy	643	GlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys	662
Db	2773	GGACACATCTGGCTTGACAGCGTTTCTTGCCAGGACATGAACCTGCTGCTGGCAATGT	2714
Qy	663	ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys	682
Db	2713	ANACCATGATGGGGAAGCATTTATGCAATCACAATGAAGATGCTGGCGTGCATGT	2654
Qy	683	SerAspAlaSerAspMetGluLeuArgLeuValGlyLysSerArgCysAlaGlyLys	702
Db	2653	TCTGATGATCAGATCTGGAGCTAAGACTTAAGAGTGGAGGACCGCCGCTGTGCGGACA	2594
Qy	703	ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn	722
Db	2593	GTGAGGTGGAGATTTCAGAGACTGTTAGGAAGAGTGTGTGACAGAGGCTGGGACTGAAA	2534
Qy	723	IleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGlu	742
Db	2533	GAAGCTGATGTGTTTGCAGGCGAGCTGGGATGTGGATCTGCACCTCAAAACATCTTATCAA	2474
Qy	743	ProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGly	762
Db	2473	GTGTACTCCAAATCCAGCAACAACACATGCTGCTTTTAAAGTAGCTGTACAGGAAAT	2414
Qy	763	GluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsn	782
Db	2413	GAACCTCTCTTTGGAGCTGCAAGAACTGGCAATGGGTGGGCTTACCTGTGTACTACTAT	2354
Qy	783	MetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMet	802
Db	2353	GAAGAAGCCAAATTTACCTGCTCAGCCACAGGAACCCAGACTGTTTGGAGGGACATT	2294
Qy	803	ProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSer	822
Db	2293	CCCTGTTCTGGACGTGTTGAAGTGAAGCATGTGTACACAGCTGGGCTCCATCTGTGATTCG	2234
Qy	823	AspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIle	842
Db	2233	GACTTCTCTCGAAGCTGCCAGCGTTCTATGCAGGAATTTACAGTGTGGCAGACATTGCT	2174
Qy	843	SerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPhe	862
Db	2173	TCTATCTCTGGGGAGCTCACCTTTGGAGAGGGAATTTGACAGATCTGGCTTGAAATTC	2114
Qy	863	GlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp	882
Db	2113	CAGTGTGAGGACATGATGCTCCATCTTTCACCTGCTGCCAGTAGCACCCGCCCAAGGA	2054
Qy	883	ThrCysIleHisSerArgGluValGlyValValCysSerArgTyThrAspValArgLeu	902
Db	2053	ACTTTAGCCACAGCAGGAGTGTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTG	1994
Qy	903	ValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGly	922
Db	1993	GTGAATGGCAACACCCGCTGTGAGGCGACAGTGGAGCTCAAAACGCTTGGTGGCTGGGA	1934
Qy	923	SerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuC	

Qy	1083	pheAsnAlaThrValserAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspAsp	1102
Db	1453	ATTAATGCCCATGGTTCGTCTCATTTTGGGGAAGAAACAGAGGCCCATCTGGCTGGATGAG	1394
Qy	1103	LeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGln	1122
Db	1393	ATGAATATGCAATGGAAAGAATCCCGCATTTGGCAGTGCCATTACACAGGCTGGGGCAG	1334
Qy	1123	HisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArg	1142
Db	1333	CAAAATTTGACGCACACAGAGAGATCGGGAGATTATCTGCTCAGAAATTCATCTCTCTGAGA	1274
Qy	1143	LeuTrpSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGly	1162
Db	1273	CTGACAGATGAAGCCACGACAGAGAGCGCTGTGCAGGGCGTCTGGAAGATTTTTCACAATGGA	1214
Qy	1163	ThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArg	1182
Db	1213	GCTTGGGGACATCTTGGCAAGAGTAGCATGTCTGAAACCAGTGTGGGTGGTGTGCAGG	1154
Qy	1183	GlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySer	1202
Db	1153	CAGCTGGGTGTGCACACAAAGGGAAATCAACCTGCATCTTTAGACAAGGCCATGTCC	1094
Qy	1203	GlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCys	1222
Db	1093	ATTCCCATGTGGTGTGACAATGTTCAAGTGTCCAAAGACCTGCACCGTGGCAGTGC	1034
Qy	1223	LeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleThrCys	1242
Db	1033	CCATCATCTCCATGGGAGAAGAGATGGCCAGCCCTCGGAGGAGACCTGGATCACATGT	974
Qy	1243	GluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrp	1262
Db	973	GACAAACAAGATAAGACTTCAGGAAGGACCCCACTTCTCTGTCGACGTGTGGAGATCTGG	914
Qy	1263	HisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGluVal	1282
Db	913	CATGAGAGTTCTCTGGGGACATGTGTGATGACTCTTGGGACTTTGGACGATGCTCAGGTG	854
Qy	1283	ValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAlaSerPheGly	1302
Db	853	GTGTGTCAACAACCTGGCTGTGGTCCAGCTTTGAAAGCATTCAAAAGACGACAGATTGGT	794
Qy	1303	GlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeu	1322
Db	793	CAGGGGACTGGACCGATATGGCTCAATGAAGTGAAGTGCAAGGGAATAGTCTTCCTTG	734
Qy	1323	TrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGly	1342
Db	733	TGGGATTTGCTCCTCCAGACGTGGGCCCATYAGTAGTGTGGGCACAGGAACAGCCTGCA	674
Qy	1343	ValArgCysSerGlyGlnSerLeuLys-----SerLeuAsnAlaSer	1356
Db	673	GTGAATTTGCACAGATATTTTCAGTGCAGAAAAACCCACAAAAAGCCACAACAGGTCGTCA	614
Qy	1357	SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPhe	1376
Db	613	TCCCGTCAGTCATCCTTTATTCAGTCGCGGATTCCTTGGGGTTGTCTGTGGCCATTTTC	554
Qy	1377	Ile-----LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArg	1394
Db	553	GTCCGATATTCTTCTTGACTAAAGCGGACGACAGACAGACGCG-----CTTGCA	503
Qy	1395	ValSerThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCys	1414
Db	502	GTTTCTCAAGAGGAGAGAACTTAGTCCCAACAAATTCATACCGGAGATCAATCTCTGC	443
Qy	1415	LeuLysArgGluAsp	1419
Db	442	CTGAATGCAGATGAT	428

REPORT OF

RESULTS

AA160436/C
ID AAT60436 standard: CDNA: 3834 BP.

AA
AC
AAI60436:

22-OCT-2001 (first entry)

Human polynucleotide SEQ I

KW Human; nootropic; immunosuppressant;
KW Human; nootropic; immunosuppressant;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.

AA
OS
Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312

PR 03-AUG-2000; 2000US=0653450.
PR 14-SEP-2000; 2000US=0662191

PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.

AA (HYSE-) HYSEQ INC. PA

Tanq YT, Liu C, Asundi V,

PI Zhao QA, Zhou P, Goodrich

DR WPI; 2001-442253/47.

XX

PT such as central nervous system injuries -
XX
Claim 1: SEQ ID NO 4425: 10078bp: English.

CC The invention relates

CC the encoded polypeptides (accession numbers) immunosuppressant and cytostatic activity.

CC of the invention may be used to treat diseases of the peripheral ne

CC localised neuropathies and central nervous

CC lateral sclerosis, and Shy-Drager Syndrome

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

SQ Sequence 3834 BP; 979 A; 1082 C; 756 G; 1017 T; 0 other;

Alignment Scores:

505

Best Local Similarity:	56.158	Mismatches:	294
Overall Match:	42.148	Indels:	12

DB:	22	Gaps:	4
US-09-759-130B-381 (1-1453) x AA160436 (1-3834)			
Qy	364	GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu	383
Db	3613	GGAAACAGCAAGGAGCTGAGGCTAGTGATGGTGAACAACAGTGTAGCGGGAGAGTGGAA	3554
Qy	384	ValArgIleHisGlnTrpThrIleCysAspGlnAsnTrpLysAsnGlnGlnAla	403
Db	3553	GTCAAGATCCAGGAGGAGTGGGAACGGTGTGTAATAATGCTGGACCATGAAGCGGTC	3494
Qy	404	LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys	423
Db	3493	TCGTGATTTGAACCACTGGGATGTCCAACGTATATCAAAAGCCCTGGATGGCGTAAT	3434
Qy	424	ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer	443
Db	3433	TCACGTCCAGGTTCTCGACGCATTTGGATGGATCATGTTCTTGCTGGGAATGAGTCA	3374
Qy	444	AlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr---CysPheArgArgSer	462
Db	3373	GCCTCTTTGGGATTGCAACATCATGATGGTGGGAAACATAGTAAGTGTACTCACCAACA	3314
Qy	463	AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis	482
Db	3313	GATGCTGGAGTGACCTGCTCAGATGGATCCCAATTTGGAAATCAGGCTGACGCGTGGAGG	3254
Qy	483	SerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHis	502
Db	3253	AAATATGTTGCTGGAAGAATAGAGATCAAAATTCCAAGGACCGTGGGAAACAGTGTGTGAT	3194
Qy	503	AspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPro	522
Db	3193	GATAACTTCAACATAGATCATGCATCTGTCAATTTCTGACAACTTGAATGTGGAAGTCT	3134
Qy	523	MetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAspAsp	542
Db	3133	GTCAGTTTTCTCTGGTTCATCTAATTTTGGAGAAAGCTCTGGACCAATCTGCTTTGATGAT	3074
Qy	543	ValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys	562
Db	3073	CTTATATGC AACGGAATGAGTCAGCTCTCTGGAAGTGC AACACATCAAGGATGGGAAG	3014
Qy	563	HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly	582
Db	3013	CATAACTGTGATCATCTCTGAGGATGCTGGAGTATTGCTCAAGGGAGCAGATCTGACG	2954
Qy	583	LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGly	602
Db	2953	CTGAGACTGTGATGATGGAGTCACTGAATGTTTCAGAAAGATTAGAAAGTGAAGTCCCAAGG	2894
Qy	603	ArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSer	622
Db	2893	GAATGGGGACAATATGTATGACGCTGGACAGTACGATGCTGCTGGCATGCAAG	2834
Qy	623	GlnLeuAspCysProSerIleLeuGlyMetGlyLeuGlyAsnAlaSerThrGlyTyr	642
Db	2833	CAACTGGGATGTCCTCACTCCGCTCACACCATTTGGTCAGATTACGCCAGTAAAGGATTT	2774
Qy	643	GlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys	662
Db	2773	GGACACATCTGGCTTGACACGGTCTTCTGCCAGGCACATGAACCTGTCTCTGCAATGT	2714
Qy	663	ArgAsnSerClyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys	682
Db	2713	AAACACCATGAATGGGGAAGCATTTATTGCAATCAATGAAGATGCTGGCGTGACATGT	2654
Qy	683	SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLys	702
Db	2653	TCGTGATGATCAGATCTGGAGCTTAGACATTAGAGTGGAGCCGCCGCTGTCTGGGACA	2594
Qy	703	ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn	722

Conservative: 161
Mismatches: 205

Query Match: 41.74% Indels: 45
DB: 22 Caps: 5

US-09-759-130B-381 (1-1453) x AA158650 (1-3811)

Qy 364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383
Db 222 GGAACAGACAGACAGCTAGTGGATGGTGAAGAAAGTGTAGCGGAGAGTGGAA 281

Qy 384 ValArgIleHisGluGlnTrpThrIleCysAspGlnAsnTrpLysAsnGluAla 403
Db 282 GTGAAGTCCAGGAGGAGTGGGAACGCTGTGAATATGCTGGAGCATGGAACGGCTC 341

Qy 404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgAlaLys 423
Db 342 TCTGTGATTGTAAACAGCTGGGATGCCAAGTCTATCAAGCCCTGGATGGCTAAT 401

Qy 424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443
Db 402 TCCAGTGCAGGTTCTGGACGCAATTTGGATGGATCATGTTCTTCTGCTGGGAATGAGTCA 461

Qy 444 AlaLeuTrpAspCysThrTyraSpGlyLysAlaLysArgThr---CysPheArgArgSer 462
Db 462 CTCTTTGGGATGCAACATGATGGATGGGAAGCATAGTAAGTCTACTCACCACAA 521

Qy 463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis 482
Db 522 GATGCTGGAGTGACCTGCTCAGATGGATCCCAATTTGGAAATGAGGCTGACCCGTTGGAGGG 581

Qy 483 SerProCysThrGlyArgLeuGluValLysTyrglnGlyGluTrpGlyThrValCysHis 502
Db 582 AATATGTTCTTGGGAAGATAGATCAATTTCCAAAGACGCTGGGAACAGTGTGTGAT 641

Qy 503 AspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPro 522
Db 642 GATAACTTCAACATAGATCATGTCATGTCATTTGTAGACAACATGAATGTGGAGTGCCT 701

Qy 523 MethisValPheGlyMetThrTyraPheLysGluAlaSerGlyProIleTrpLeuAspAsp 542
Db 702 GTCAGTTTCTGCTGTCATTAATTTTGGAGAAGGCTCTGGACCAATCTGCTTTGATCAT 761

Qy 543 ValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys 562
Db 762 CTTATATCAACGGAATAGCTAGCTCTCTGGAAGTCAACATCAAGAGTGGGGAAG 821

Qy 563 HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly 582
Db 822 CATAACTGTGATCATGCTGGAGTGTCTGGAGTGTCTCAAGGGAGCAGATCTGAGC 881

Qy 583 LeuArgLeuValGlyLysSerAsnArgCysSerGlyArgLeuGluValTyraPheGlnGly 602
Db 882 CTGAGACTGGTAGATGGAGTCAATGTTTCAGGAAGATAGAGTGGAGTTCAGGA 941

Qy 603 ArgTrpGlyThrValCysAspGlyTrpAsnSerLysAlaAlaValValCysSer 622
Db 942 GAATGGGGACAATATGATGACGGCTGGGACAGTACGATGCTGCTGGCGATGCAAG 1001

Qy 623 GlnLeuAspCysProSerSerIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyra 642
Db 1002 CAACTGGGATGTCACACTGCGTCACACCAATGCTGAGTTAACCCAGTAAGGGATTT 1061

Qy 643 GlyLysIleTrpLeuAspValSerCysAspGlyAspGlySerAspLeuTrpSerCys 662
Db 1062 GCACACATCTGGCTTGACAGCGTTTCTTCCAGGGACATGAACCTGCTGTGCGCAATGT 1121

Qy 663 ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys 682
Db 1122 AAAACACCATGATGGGAAACATATTATTCGAATCAATCAATGAAGTGTGGCGCTGACATGT 1181

Qy 683 SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerArgCysAlaGlyLys 702
Db 1182 TCTGATGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGCGACCCGCTGTGCTGGGACA 1241

Qy 703 ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn 722
Db 1242 GTTCAGGTGGAGATTACAGACTCTTAGGGAAGTGTGTGACAGAGGCTGGGACTGAAA 1301

Qy 723 IleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGlu 742
Db 1302 GAAAGCTGATGTGGTTTCCAGCAGCTGGGATGTGATCTGCACATAAACAATCTATCAA 1361

Qy 743 ProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGly 762
Db 1362 GTGTACTCCAAAATCCAGGCAACACACATGCTGTTCTTAAGTAGCTGTAAACGAAT 1421

Qy 763 GluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsn 782
Db 1422 GAAACTTCTCTTTGGGACTGCAAGAACTGCAATGGGGTGGACTTACCTGTGATCATTAT 1481

Qy 783 MetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMet 802
Db 1482 GAAGAAGCCAAAATTTACCTGCTCAGCCCAACAGGAACCCAGCTGGTTGGAGGGACATT 1541

Qy 803 ProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSer 822
Db 1542 CCTGTCTTGGACGTGTGAAGTGAAGCATGGTGACAGCTGGGCTCCATCTGTGATTTCG 1601

Qy 823 AspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIle 842
Db 1602 GACTTCTCTGGAAGCTGCCAGCTTCTATGACGGAATTTACAGTGTGGCACACTTGTCTC 1661

Qy 843 SerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPhe 862
Db 1662 TCTATCTCTGGGGAGCTCACTTTGGAGAGGAAATGACAGATCTGGGCTGAAGAAATTC 1721

Qy 863 GlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp 882
Db 1722 CAGTGTGAGGACATGATGCTCCATCTTCTCACTCTGCCAGTAGCACCCGCCCAAGA 1781

Qy 883 ThrCysIleHisSerArgGluValGlyValValCysSerArg--- 896
Db 1782 ACTTGTAGCCACAGCAGGGATGTTGGAGTAGTCTCTCAAGTAAGACCCACAGAAAACATCT 1841

Qy 896 --- 896

Db 1842 TTAATTTGTTCTTACTGTGAAAGGGACAGGTTAGGGAGTCAATAGCTGTCTTTTCTTA 1901

Qy 897 --- TyrThrAspValArgLeuValAsnGlyLysSerGlnCys 909
Db 1902 AAGCCCTCTCTCTCCAGGATACACAGAAATTCCTGTTGTAATGGCAGACCCCGTGT 1961

Qy 910 AspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTrp 929
Db 1962 GAGGACAGTGGAGCTCAAAACGCTTGGTCTGGGATCCCTCTGTAACCTCTCACATGG 2021

Qy 930 AspProGluAspAlaArgValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThr 949
Db 2022 GACATAGAGATGCCCATTGTTTCCAGCAGCTTAATATGGAGTGGCTTCTCTTCTACC 2081

Qy 950 ThrGlyGlyLysTyraIleGlyLysArgSerValArgValTrpGlyHisArgPheHisCys 969
Db 2082 CCAGGAGGACGCGCTTTTGGAAAGGAAATGCTCAGATCTGGAGGCATATGTTTCACTGC 2141

Qy 970 LeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThrValLeuGlyAlaProCys 989
Db 2142 ACTGGAGTACGACACATGGGAGATTGCTCTGTAACATGCTCTAGGTCTTCATTATGT 2201

Qy 990 IleHisGlyAsnThrValSerValIleCysThrGlySerLeuThrGlnProLeuPhePro 1009
Db 2202 CCTTCAGAGCAAGTGGCTCTGTATCTGTCAGGAACCACTCCCAACACTGTCTCGG 2261

Qy 1010 CysLeuAlaAsnValSerAspProTyraLeuSerAlaValProGluGlySerAlaLeuIle 1029
Db 2262 TGCATATCATCTTTGGGCCCCAACAGAGCTACCATTCAGAGAAAGTGTCTGTGCC 2321

Qy 1030 CysLeuGluAspLysArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgVal 1049

Db 2322 TGCATAGAGAGTGGTCAACTTCGCTGTAATGAGGAGGTGCTGCTGGGAGATA 2381
Qy 1050 GluIleThrHisAspGlyPheThrGlyThrIleCysAspAspGlyTrpAspLeuSerAsp 1069
Db 2382 GAGATCTATCATGAGGAGTCCGCGGACCATCTGCTGATGAGCTGGGACCTGAGTGAT 2441
Qy 1070 AlaHisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAla 1089
Db 2442 GCCACAGTGGTTGGACAGAGCTGGGCTGTGGAGAGGCCATTAATGCCACTGGTCTGCT 2501
Qy 1090 HisPheGlyGluCysGlySerGlyProIleThrLeuAspAspLeuAsnCysThrGlyThrGlu 1109
Db 2502 CATTTTGGGAAAGAACAGGCGCCATCTGCTGGATGAGTGAATGCAATGGAAAGAA 2561
Qy 1110 SerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGlu 1129
Db 2562 TCCCGCATTTGGCAGTGGCAATTCACACGCTGGGGGACCAAAATGGAGGCAAGAG 2621
Qy 1130 AspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGluThr 1149
Db 2622 GATCGGGAGTTATCTGCTCAGCAATTCATGCTGAGACTGACCAAGTGAAGCCAGCAGA 2681
Qy 1150 GluSerCysAlaGlyArgLeuGluValPheThrAsnGlyThrTrpGlySerValGlyArg 1169
Db 2682 GAGGCTGTGCAAGGCGTCTGGAAGTTTTTACAAATGGAGCTTGGGGCACTTTGGCAAG 2741
Qy 1170 ArgAsnIleThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsn 1189
Db 2742 AGTAGCATGCTGAAACCACTGTGGTGTGGTGTGTCAGGCACCTGGGCTGCACACAAA 2801
Qy 1190 GlyValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAsp 1209
Db 2802 GGGAAATCAACCTGCATCTTTAGCAAGGCCATGTCCATTCCTCCATGGGTGGACAAT 2861
Qy 1210 IleGlnCysProLysThrHisIleSerIleThrGlnCysLeuSerAlaProTrpGluArg 1229
Db 2862 GTTCAGTGTCCAAAGAGGACCTGACAGCTGTGGCAGTGGCCATCATCTCCATGGGAGAG 2921
Qy 1230 ArgIleSerProAlaGluThrTrpIleThrCysGluAspArgIleArgValArg 1249
Db 2922 AGATGCCAGCCCTCGGAGAGACCTGGATCATCATGTGCAACAAGATAGACTTCAG 2981
Qy 1250 GlyGlyAspThrGluCysSerGlyArgValGluIleThrPheAlaGlySerTrpGlyThr 1269
Db 2982 GAAGGACCCATCTCTGTCTGGAGCTGTGGAGATCTGCATGGAGTTCTCTGGGGGACA 3041
Qy 1270 ValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCys 1289
Db 3042 GTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTGTGTGTCAACAACCTTGGCTGT 3101
Qy 1290 GlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrp 1309
Db 3102 GTGCCAGCTTGAAGACCATCAAGAGCAGAGTGTGGTCAGGGGACTGGCCGATATGG 3161
Qy 1310 LeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysPro 1329
Db 3162 CTCATGAAAGTGAAGTGAAGGGAATGAGTCTCTTGTGGGATGTCTCTCCCGACAGCC 3221
Qy 1330 TrpGlyLysSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSer 1349
Db 3222 TGGGGCCATAGTGTGGGACAGAGGAGCGTGCAGTGCAGTGAATGGACAGATATTTCA 3281
Qy 1350 LeuLys-----SerLeuAsnAlaSerSerGlyHisLeuAlaLeuIle 1363
Db 3282 GTGCAGAAAAACCCCAAAAGCCACAGAGTGGCTCATCCGTCAGTCATCTTAT 3341
Qy 1364 LeuSerSerIlePheGlyLeuLeuLeuValLeuPheIle-----LeuPheLeuThr 1381
Db 3342 GCAGTCGGGATCTTGGGGTGTCTGTGGCCATTTTTCGTCGATTTATCTTCTGACT 3401
Qy 1382 TrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgGly 1401

Db 3402 AAAAGCGAAGACAGACAGACAGCGG-----CTTGCACTTCTCAAGAGGAGAGAAC 3452
Qy 1402 SerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAsp 1419
Db 3453 TTAGTCCACAAATTCATACCGGAGATGAATCTTGCCTGAATGCAGATGAT 3506
RESULT 10
AAS94922
ID AAS94922 standard; DNA; 3707 BP.
XX AAS94922;
XX 14-FEB-2002 (first entry)
XX Human DNA sequence #177 expressed during foam cell differentiation.
XX Human; foam cell differentiation; atherosclerosis; cerebral stroke;
KW cardiovascular disorder; coronary artery disease; gene therapy; ds.
XX Homo sapiens.
XX WO200177389-A2.
XX 18-OCT-2001.
XX 04-APR-2001; 2001WO-US11128.
XX 05-APR-2000; 2000US-195106P.
XX (INCY-) INCYTE GENOMICS INC.
XX Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;
PI Tai J;
XX WPI; 2002-010925/01.
XX Composition useful for diagnosis of conditions, disorders or diseases
PT associated with atherosclerosis, comprises several polynucleotides that
PT are differentially expressed in foam cell development -
XX Claim 1; Page 232-233; 315pp; English.
XX The present invention relates to the isolation of human polynucleotide
CC sequences that are differentially expressed during foam cell
CC differentiation. The polynucleotide sequences of the invention or a
CC composition comprising these polynucleotides are useful as a high
CC throughput method for detecting altered expression of one or more
CC polynucleotides in a sample. The polynucleotides can be used in the
CC diagnosis of disorders associated with foam cell development such as
CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as
CC coronary artery disease. The polynucleotide sequences can also be used
CC as PCR primers and probes. The polynucleotides of the invention are also
CC useful in gene therapy. AAS94746-AAS95021 represent the human
CC polynucleotide sequences of the invention which are differentially
CC expressed during foam cell differentiation.
XX Sequence 3707 BP; 984 A; 731 C; 1062 G; 930 T; 0 other;
SQ Alignment Scores:
Pred. No.: 2,61e-207 Length: 3707
Score: 3377.00 Matches: 598
Percent Similarity: 71.00% Conservative: 161
Best Local Similarity: 55.94% Mismatches: 294
Query Match: 41.50% Indels: 16
DB: 24 Gaps: 4
US-09-759-130B-381 (1-1453) x AAS94922 (1-3707)

Qy 364 GlyAlaAspLeuGluArgLeuAlaAspGlySerAsnAsnCysSer-GlyArgValG 383
Db 222 GGAACAGACAGAGAGCTGAGGCTAGTGGATGTGAAACAAAGTGTACCGGGAGAGTGA 281
Qy 383 uValArgIleHisGluGlnTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAl 403

Db	282	AGTGAAGTCCAGGAGAGTGGGAACGGTGTGTAATAATGGCTGGAGCATGGGAAGCGGT	341
Qy	403	aLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys	423
Db	342	CTCTGTGATTGTGAACCAAGTGGGATGTCCTCAACTGCTATCAAGCCCTGGATGGGTAA	401
Qy	423	sProSerAsnGluAlaAraGspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSe	443
Db	402	TTCACGTGCAGGTCTTCGGACGATTTGGATGCATCATCTCTCTCTCGTGGAAATGAGTC	461
Qy	443	rAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr--CysPheArgArgSe	462
Db	462	AGCTCTTTGGGATTGCAACATGATGGATGGGAAGCATAGTAACGTGTACTCCACCAACA	521
Qy	462	rAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHi	482
Db	522	AGATGCTGGAGTCACCTGCTCAGATGGATCCCAATTTGGAATGAGGCTGACGCGTGGAGG	581
Qy	482	sSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHi	502
Db	582	GAATATGTCTTCTGGGAAGATAGACATCAAAATTCCAAGGACGGTGGGAACAGTGTCTGA	641
Qy	502	sAspAlaGlyTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPr	522
Db	642	TGATACTTCAACATAGATAGCATCTGTCATTTGTAGACAACCTGAATGTGGAAGTGC	701
Qy	522	oMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAspAs	542
Db	702	TGTCAGTTCTCTGGTTCATCTAATTTGGAGAAGCTCTGGACCAATCTGCTTTGATGA	761
Qy	542	pValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLy	562
Db	762	TCTATATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAAACATCAAGGATGGGAAA	821
Qy	562	sHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGl	582
Db	822	GCATAACTGTGATCATGCTGGAGATGCTGGAGTATTGCTCAAGGGAGCAGATCTCGAG	881
Qy	582	yLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGl	602
Db	882	CCTGAGACTGGTAGTGGAGTCACTCAATGTTTCAGGAAGATTAGNAGTGAGATTCRAAG	941
Qy	602	yArgTrpGlyThrValCysAspGlyTrpAsnSerLysAlaAlaValValCysSe	622
Db	942	AGAATGGGGACAATATGTATGACGGCTGGCAGACTTACGATGCTGCTGTGGCATGCAA	1001
Qy	622	rGlnLeuAspCysProSerSerIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTy	642
Db	1002	GCAACTGGGATGTCCAACTCGCGTCACGCCAATGCTCGAGTTAACGCCAGTAAGGATT	1061
Qy	642	rGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCy	662
Db	1062	TGACACATCTGCTTGCACAGCGTTCTTTCGCAGGACATGAACCTGCTGCTGGCAATG	1121
Qy	662	sArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCy	682
Db	1122	TAACACCATGAATGGGGAACATATTGTGCAATCACAATGAAGATGCTGCGTGCACATG	1181
Qy	682	sSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLy	702
Db	1182	TTCTGTATGATCATGATCTGCAGCTAAGACTTATAGAGCTGGAGCGACCGCTGCTGGAC	1241
Qy	702	sValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAs	722
Db	1242	AGTTGGGTGGAGATTACAGAGACTGTGAGGAAGTGTGTGACAGAGCTGGGACGTAA	1301
Qy	722	nIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGl	742
Db	1302	AGAACTGATGTGTTTGCAGCAGCTGGGATGTGGATCTGCATCTCAAAACATCTTATCA	1361
Qy	742	uProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGl	762

D	b	1362	AGTGACTCCAAATCCAGCAACAACACATGGCTGTTTCTTAAGTACTGCTGTACCGAAA	1421
Q	y	762	yGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAs	782
D	b	1422	IGAAACTTCCTTTGGGACTGCAAGAACTGGCAATGGGTGGACCTTACCTGTGATCACTA	1481
Q	y	782	nMetGluAlaSerLeuIleCysSer-AlaHisArg-GlnProArgLeuValGlyAlaAsp	801
D	b	1482	TGAAGAAGCCAAATTAACCTGCTCAGGCCCCACAGGGGAACCCACAGACTGGTTGGAGGGAC	1541
Q	y	802	MetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAsp	821
D	b	1542	ATTCCCTGTTCTGGACGCTGTTGAAGTGAACATGGTGACACGCTGGGCTCCATCTGTGAT	1601
Q	y	822	SerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAla	841
D	b	1602	TCAGACTTCCTCGGAAGCTGCCAGCGTCTATGCAGGGAATTACAGTGGGCACAGTT	1661
Q	y	842	IleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLys	861
D	b	1662	GTCTCTATCTCGGGGGAGCGTCACTTTGGAGAGGGAATGGACAGATCTGGGCTGAAGA	1721
Q	y	862	PheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGlu	881
D	b	1722	TTCCAGTGTGAGGACATGATGCCATCTTTCACCTGCCCAGTAGCAGCCCCGCCAGAA	1781
Q	y	882	AspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspValArg	901
D	b	1782	GGAATTGTAGCCACAGCAGGAGTGTGGAGTAGTCTGCTCAAGATACACAGAAATTCGC	1841
Q	y	902	LeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrp	921
D	b	1842	TTGTGTAATGGCAAGACCCGCTGTGAGGCAGAGTGGAGCTCAAAACGCTTGTGGTCCCTGG	1901
Q	y	922	GlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeu	941
D	b	1902	GGATCCCTCTGTAACTCTCACTGGGACATAGAAGATGCCATGTTCTTTGCCAGCAGCTT	1961
Q	y	942	SerCysGlyThrAlaLeuSerThrGlyGlyLysTyrIleGlyGluArgSerValArg	961
D	b	1962	AAATGTGAGTTCGCCCTTCTACCCAGGAGGAGCAGCGTTTGGAAAAGAAATGGTCAG	2021
Q	y	962	ValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMet	981
D	b	2022	ATCTGGAGGCATATGTTTCACTGCATCTGGGACTGGAGCAGACATGGGAGATTGTCTCTA	2081
Q	y	982	ThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGly	1001
D	b	2082	ACTGCTCTAGTGTCTCATATATGCTCTCAGACAGTGGCCCTCTGTAACTGTCTCAGGA	2141
Q	y	1002	SerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAla	1021
D	b	2142	AACCAAGTCCCAACACTGCTCTGCTGCATTTATCGTCTTTGGGCCCAACAGGCCCTACC	2201
Q	y	1022	ValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGly	1041
D	b	2202	ATTCACAAGAAAGTGTCTGGCCCTGCATAGAGAGTGGTCAACTTCGCCTGGTAAATGGA	2261
Q	y	1042	AspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCys	1061
D	b	2262	GGAGGTCCGCTGCTCGGGAGAGTAGAGATCTATCATGAGGGCTCCTGGGGCACCACATCTGT	2321
Q	y	1062	AspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyVal	1081
D	b	2322	GATGACAGCTGGGACCTTGAGTGAATGCCACGTGGTTTGCAGACAGCTGGGCTGTGGAGAG	2381
Q	y	1082	AlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAsp	1101
D	b	2382	CCCATTAATGCCACTGGTTCTGCTCATTTTGGGAAGGACAGAGGCCCATCTGGCTGGAT	2441
Q	y	1102	AspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGly	1121
D	b	2442	GAGATGAATGCAATGGAAAGAAATCCCGCATTTTGGCACTGGCACTTACACGGCTGGGG	2501

QY 1122 GlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeu 1141
 DB 2502 CAGCAAAATTCAGGACACAGAGATCGGGAGTTATCTGTCAGATTCTATCTCTG 2561
 QY 1142 ArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsn 1161
 DB 2562 AGACTGACCAAGTGAAGCCAGCAGAGAGCGCTCTGAGGCGCTCTGGAAGTTTTCACAAAT 2621
 QY 1162 GlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCys 1181
 DB 2622 GGAGCTTGGGCACTGTGGCAAGAGTAGCATGTCTGAACACACTGTGGGTCTGTGTGTC 2681
 QY 1182 ArgGlnLeuGlyCysGlyGluValValSerLeuAlaProLeuSerLysThrGly 1201
 DB 2682 AGGCAGCTGGGTGTGCACAAAGGAAATCAACCTGTCATCTTAGACAAAGCCATG 2741
 QY 1202 SerGlyPheMetTrpValAspPileGlnCysProLysThrHisIleSerIleTrp-Gl 1221
 DB 2742 TCATTCCCATGTGGGTGGACAATGTTCAAGTGTCCAAAAGGACCTGACACAGCTGTGGCA 2801
 QY 1221 nCysLeuSerAlaProTrpGluArgGlySerSerProAlaGluGluThrTrpIleTh 1241
 DB 2802 GTGCCATCATCTCCATGGGAGAGAGACTGGCCAGCCCTCGAGGAGACCTGGATCAC 2861
 QY 1241 rCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluI 1261
 DB 2862 ATGTGACACAGATAAGACTTCAGGAAGACCCACTTCTGTTCTGGAGCTGTGGAGAT 2921
 QY 1261 eTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGl 1281
 DB 2922 CTGGCATGGAGGTCTCTGGGGACAGTGTGTATGACTCTTGGGACTTGGAGCATGCTCA 2981
 QY 1281 uValValCysGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPh 1301
 DB 2982 GGTGGTGTGTCACAACTTGGCTGTGTGTCAGCTTTGAAGCATTCAAAGACAGAGTT 3041
 QY 1301 eGlyGlnGlyThrGlyThrIleTrpLeuAspMetArgCysLysGlyAsnGluSerPh 1321
 DB 3042 TGCTCAGGGGACTGGACCATATGGCTCAATGAAGTGAAGTGAAGGGAATGAGTCTTC 3101
 QY 1321 eLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAl 1341
 DB 3102 CTTGTGGGATGTCCTGCCACAGCTGGGGCCATAGTGTAGTGTGGGCAAGGAAGACGC 3161
 QY 1341 aGlyValArgCysSerGlyGlnSerLeuLys-----SerLeuAsnAl 1355
 DB 3162 TGCAAGTGAATGGACAGATATTTTCAGTGCGAGAAACCCACAAAGCCACACAGTGC 3221
 QY 1355 aSerSerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLe 1375
 DB 3222 CTCATCCCGTCAGTCATCTTATTCAGTCGCGGATCTCTGGGTGTTCTGTGGCCAT 3281
 QY 1375 uPheIle-----LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLe 1393
 DB 3282 TTTCTGTCGCATATTCTTCTGTGATATAAAGCGAAGACAGACAGACGCG-----CT 3332
 QY 1393 uArgValSerThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluTh 1413
 DB 3333 TGCAGTTTCTCAGAGGAGAGACTTAGTTCACCAAAATTCATACCGGAGATGAATTC 3392
 QY 1413 rCysLeuLysArgGluAsp 1419
 DB 3393 TTGCCTGAATGACATGAT 3411
 RESULT 11
 AAF45141
 ID AAF45141 standard; cDNA; 4308 BP.
 AC
 AC AAF45141;
 XX
 XX
 XX
 DT 30-MAR-2001 (first entry)
 XX

DE XX Bovine WC1 ORF.
 KW Gene therapy: TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 KW neuropsychiatric; psychoactive substance use; anxiety; ss.
 OS Bos sp.
 XX WO200077239-A2.
 PN 21-DEC-2000.
 PD 24-MAY-2000; 2000WO-US14858.
 PF 14-JUN-1999; 99US-0333159.
 PR (MILL-) MILLENNIUM PHARM INC.
 PA McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
 PI WPI; 2001-032313/04.
 DR P-P5DB; AAB66088.
 DR TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -
 XX Disclosure: Fig 2; 359pp; English.
 XX The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder. The present sequence is a
 CC sequence used in a sequence homology alignment with the TANGO/INTERCEPT
 CC sequences of the present invention.
 XX SQ Sequence 4308 BP; 866 A; 1093 C; 1410 G; 939 T; 0 other:
 Alignment Scores:
 Pred. No.: 2,74e-186 Length: 4308
 Score: 3082.00 Matches: 613
 Percent Similarity: 55.85% Conservative: 208
 Best Local Similarity: 41.70% Mismatches: 495
 Query Match: 37.87% Indels: 154
 DB: 22 Gaps: 23
 US-09-759-130B-381 (1-1453) x AAF45141 (1-4308)
 QY 28 CysIleLeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGlyThrAspLeuGlu 47
 DB 37 TGTGTCTCTCTCTCTCGGC-----ACCATGGTGGTGGTCAAGCTCTGGAG 81
 QY 48 LeuArgLeuValAsnGlyAspGlyProCysSerGlyThrValGluValLysPheGlnGly 67
 DB 82 CTCGAGTTGAAGGATGGAGTCCATCGCTGTGAGGGAGAGTGAAGTGAAGACCAAGGA 141
 QY 68 GlnTrpGlyThrValCysAspAspGlyTrpAsnThrThrAlaSerThrValValCysLys 87
 DB 142 GAATGGGGCACAGTGGATGGTTACAGTGGACATGAAGGATCATCTGTAGTGTGAGGA 201
 QY 88 GlnLeuGlyCysProPheSerPheAlaMet-----PheArgPheGlyGlnAlaVal 104

Db 202 CAGCTGGGGTGGAGCTGCCATTTGCTTCTGGAGGGGCTTATTGCGCCAGGACTT 261
Qy 105 ThrArgHisGlyLysIleTrpLeuAspValSerCysTyrGlyAsnGluSerAlaLeu 124
Db 262 -----GGCCCATTTGGCTTTTGTATCTTCATCTGTGAGGAGCAGAGTCAACTGTC 312
Qy 125 TrpGluCysGlnHis-----ArgGluTrpGlySerHisAsnCysTyrHisGlyGlu 141
Db 313 AGTGACTGTGAGCATTTCTAATATTAAAGACTATCGTAATGATGCTATAATCATGGTCGG 372
Qy 142 AspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsn 161
Db 373 GATGCTGGAGTAGTCTGCTCAGGA-----TTTGTGCGTCTGGCTGGAGGGCAT 420
Qy 162 AsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAsp 181
Db 421 GGACCTGTCTCAGGCGCAGTAGAAGTGCATCTCGGAAGCTTGGATCCCAAGTCTGTGAT 480
Qy 182 AspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSerSer 201
Db 481 GGGAACTTCACATCTGCACCTGCCAGATCATCTGTGCAGAGTTGGTGTGGCAAGGCT 540
Qy 202 PheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAspAsp 221
Db 541 GTGTCTGTCTGGGACATGAGCTCTTCAGAGAGTCCAGTGCCAGGCTCGGGCTCGAAGAG 600
Qy 222 IleLeuCysGlnGlyAsnGluAlaLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsn 241
Db 601 TTCAGGTGTGAGGGGGAGGAGCCCTGAGCTCTGGGTCTGCCAGGAGTCCCTGTCCAGGG 660
Qy 242 HisAspCysSerHisAsnGluAspValThrLeuThrCystyrAspSerSerAspLeuGlu 261
Db 661 GGCACGTGTACACACAGTGGATCTGCTCAGGTGTTGTTTTCAGCATACTCAGAAGTCCGG 720
Qy 262 LeuArgTrpValGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGly 281
Db 721 CTC---ATGCAAAACGGCTCTCTCAGTGTGAAGGCGAGGTGGAGATGAACATTCTTCTGGA 777
Qy 282 ArgTrpGlyThrValCysHisLysTrpAsnAsnAlaAlaAspValValCysLys 301
Db 778 CAATGGAGAGCGCTGTGTGCTCCCTGAGCTGGAGCTGCCAATGCCAATGTTATCTGCTGT 837
Qy 302 GlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGlySer 321
Db 838 CAGCTCGGCTGTGGAGTTGCCATCTCCACCCCGGAGGACCACACTTGGTGGAGAAGGT 897
Qy 322 AspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCys 341
Db 898 GATCAGATCTCTAACAGCCCGGATTTCACTGTCTGGGGCTGAGTCTCTCTGAGTGTGT 957
Qy 342 ArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCys 361
Db 958 CCTGTGACTGCCCTGGGTGGTCTGACTGTGCCATGGCAACACAGCCTCTGTGATCTGC 1017
Qy 362 SerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArg 381
Db 1018 TCAGGAAC----- 1026
Qy 382 ValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGlu 401
Db 1027 ---CAGATCAGGCTGTCCCCAG----- 1047
Qy 402 GlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArg 421
Db 1048 -----TGCAACCACTCCGTG----- 1062
Qy 422 AlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsn 441
Db 1063 TCTCAACCTACAGGCTCTGCGGCGCTCA----- 1089
Qy 442 GluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArgArg 461

Db 1089 ----- 1089
Qy 462 SerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAla 481
Db 1090 GAGACAGCGCCCTACTACTGCTCAGACAGCAG-----CAGCTCCGCGCTGGTGACGGG 1143
Qy 482 HisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCys 501
Db 1144 GCGCGTCCCTCGCGCGGAGAGTGGAGATCTTGACACAGGCTCTCTGGGCGCACCATCTGT 1203
Qy 502 HisAspArgTrpSerThrArgAsnAlaValValCysLysGlnLeuGlyCysGlyLys 521
Db 1204 GATCAGCGCTGGACCTGGACGATCCCGCGCTGTGTGTGCAGCAGCTGGGTGTGGAGAA 1263
Qy 522 ProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAsp 541
Db 1264 GCCCTCAATGCCAGGGGTCTGCTCACTTCGGGGCAGGATCAGGCCCATCTGGTTGGAC 1323
Qy 542 AspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGly 561
Db 1324 AACTTGAACTCACAGGAAGAGTCCACGCTGTGAGGTGCCCTTCCCGGGCTGGGG 1383
Qy 562 LysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrp 581
Db 1384 CAGCACAACTGCACACACAGCAGCGGGGTCTACTGTCTCA-----GAGTTCCTG 1437
Qy 582 GlyLeuArgLeuValGlyLysArgCysSerGlyArgLeuGluValValTrpPheGln 601
Db 1438 GCCCTCAGATGCTGAGTGAGGACGACGAGTCTGCTGGGTGGTGAATTTTCTACAA 1497
Qy 602 GlyArgTrpGlyThrValCysAspGlyTrpAsnSerLysAlaAlaValValCys 621
Db 1498 GGGACCTGGGCGAGTGTGCTGCCGTAACCCATGGAAGACATCACTGTCTCCACGATCTGC 1557
Qy 622 SerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr 640
Db 1558 AGACAGCTGTGCTGGGACAGT-----GGAACCCCTCAACTCTTCTGTCTGTCT 1605
Qy 641 -----GlyTrpGlyLysIleTrpLeuAspValSerCysAspGlyAspGluSer 657
Db 1606 CTTAGAGAGGTTTATAGCCACAGTGGGTGATAGATCCAGTGTCCGAAACACTGCACAC 1665
Qy 658 AspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAspCysSerHisSerGlyAsp 677
Db 1666 TCTCTCTGGCAGTCTCTCTCTGACCTTGGAAATTACAACATCATGCTCTCCAAGAGGAA 1725
Qy 678 ValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyLysSerSer 697
Db 1726 GCCTATATCTGCTGTGCAGACAGC-----AGACAGATCCCGCTGGTGGAGGTGGT 1779
Qy 698 ArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsn 717
Db 1780 CGCTGCTCTGGAGAGTGGAGATCTTCACACAGGCTCTCTGGGCGACCATCTGTGATGAC 1839
Qy 718 GlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGlyCysGlySerAlaIle 737
Db 1840 CGCTGGGACCTGGAGATGCTCCCGTGTGTGTGCAAGCAGCTGGGTGTGGAGAAGCCCTG 1899
Qy 738 ArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSer 757
Db 1900 GACCCCATGCTCTCTCTCTCTCGGACGGGATCAGGGCCCATCTGGCTGGATGAAGTG 1959
Qy 758 GlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThr 777
Db 1960 AACTGCAGAGGAGAGGATCCCAAGTATGGAGGTGCCCTTCTGGGATGGCGGCAACAC 2019
Qy 778 AlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeu 797
Db 2020 AACTGCAATCATCAAGAAGATGCAGGAGTCACTGCTCAGGATTT-----GTGCGTCTG 2073
Qy 798 ValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArg 817
Db 2074 GCTGGAGGAGATGACCCCTGCTCAGGCGGAGTAGAAGTGCAATTTCTGGAGAGACCTGGACC 2133

QY 818 SerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsn 837
 DB 2134 CCAGTGTCTGATGGAACTTACACTCCCACTGCGCCAGCTCATCTGTGCAGAGCTGGGA 2193
 QY 838 CysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThr 857
 DB 2194 TGTGGCAAGCTGTGTCTCTCTGGACACATGCCATTCAGAGATCCAGTGGCCAGGTC 2253
 QY 858 TrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleVal 877
 DB 2254 TGGGCTGAAGATTGAGTGTGATGGGGGAGCCTGAGCTCTGGTCTGCCAGAGTG 2313
 QY 878 GlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgThr 897
 DB 2314 CCTGTGCCAGGACATGTCTCCACATGGAGTGTCTCAGTGTCTGTTCAGTGTAC 2373
 QY 898 ThrAspValArgLeuVal---AsnGlyLysSerGlnCysAspGlyGlnValGluIleAsn 916
 DB 2374 ACAGAACTCCAGCTATGAAACAGGCACCTCTCAATGTGAGGGGAGGTGGAGATGAAG 2433
 QY 917 ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal 936
 DB 2434 ATCTCTGGAGATGGAGACCGCTCTGTGCTCCCACTGGAGTCTGCCCAATGCCAATGT 2493
 QY 937 LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysThrIleGly 956
 DB 2494 GTCTGTGTCAGTCTGGCTGTGGAGTGGCCATCTCCACCCAGAGGACACACACTTGGTG 2553
 QY 957 GluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu 976
 DB 2554 GRAGGAGTGTATGATATCAACAGCCCAATTTCACTGTCTAGGGCTGAGCTTCCTCC 2613
 QY 977 AspAsnCysGlnMetThrValLeuGlyAlaProCysIleHisGlyAsnThrValSer 996
 DB 2614 TGGAGTGTCTGTGACTGCTTGGTGGGCTGTGCTTCCATGGCAACACAGCCTCT 2673
 QY 997 ValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAsp 1016
 DB 2674 GTGATCTGTCAAGAAACACACCCAGGTGTGCCAGTGCACAGCTTCTGTCTCAA 2733
 QY 1017 ProTyrrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeu 1036
 DB 2734 CTTGAGGCTCTCGGCCCTCAGGAGAGTGTCTCCCTACTGTCTCAGACAGCAGGCTC 2793
 QY 1037 ArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleThrHisAspGlyPhe 1056
 DB 2794 CGCTGTGGACGGGGCGTCCCTCGCGGGGAGAGTGGAGATCTTGCACAGGCGCTCC 2853
 QY 1057 TrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLys 1076
 DB 2854 TGGGGCACCATCTGTGATGATGACTGGGACCTGGAGCTGCGGCTGTGTGTGCGAGCAG 2913
 QY 1077 LeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGly 1096
 DB 2914 CTGGGCTGTGGAGAGCCCTCAATGCCAGGGGTCTGCTCACTTCGGGGCAGGATCAGGG 2973
 QY 1097 ProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysPro 1116
 DB 2974 CCTATCTGGCTGGACACCTGAATCCACAGGAAAGAGTCCACCTGTGGAGGTGCCCT 3033
 QY 1117 SerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSer 1136
 DB 3034 TCCCGGGGCTGGGGGGCGGACGACTCGACACACAGAGGAGCGCGGGGTCTCTCTCTCA 3093
 QY 1137 GluPheThrAlaLeuArgLeuTyrrSerGluThrGluThrGluSerCysAlaGlyArgLeu 1156
 DB 3094 GAGTCTCTGGCCCTCAGGATGTGGAGC-----GAGGACCCAGAGTGTCTGGGTGGCTG 3147
 QY 1157 GluValPheTyrrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIle 1176
 DB 3148 GAGTTTCTACAAACGGGACCTGGGGCAGTGTCTGCCGAGCCCATGGAGATATCACT 3207

QY 1177 AlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaPro 1196
 DB 3208 GTGTCGGTGATGTGCAGACAGCTGGATGTGGGACAGTGAAGTCTCAACACCTCTGTT 3267
 QY 1197 LeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHis 1216
 DB 3268 GTCTCTCAGGAGGTTCTAGACCCCGGTGGTAGATTTAATTCAGTGTGCGAATAATGGAT 3327
 QY 1217 IleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGlu 1236
 DB 3328 ACCTCTCTCTGTCAGTGTCTTCTGCCCCATGGAATACAGTTCATGCTCTCCAAAGGAG 3387
 QY 1237 GluThrTrpIleThrCysGlu----- 1243
 DB 3388 GAAGCTACATCTCATGTGAAGGAAGAACCCAGAGCTGCCAACTCTCCGCCCTGC 3447
 QY 1244 -----AspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260
 DB 3448 ACAGACAGACAGAGCTCCCTCAGGGAGGAGAGACAGGAGTCTCAGGGCGGTGGAG 3507
 QY 1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280
 DB 3508 GTGTGGCACACCGCTCTCTGGGCGACCGTGTGCGATGACTCTCTGGAGCCTGGCAGAGCT 3567
 QY 1281 GluValValCysGlnClnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgaspAlaSer 1300
 DB 3568 GAGGTGTGTGTCAGCAGCTGGCTGTGGCCAGGCTGTGGAGCGCTGGAGCTGTGCAGCA 3627
 QY 1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320
 DB 3628 TTTGGCCCTGGAATGGAGCATCTGGCTGGAGAGTGCAGTGCAGGGCGCGGGAGTCC 3687
 QY 1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340
 DB 3688 TCCTCTGGGAGTGTCTTGGGAGCCCTGGGCGCAGAGGAGTGTGAAGCAGCAGGAGGAT 3747
 QY 1341 AlaGlyValArgCysSerGly----- 1347
 DB 3748 GTGTGTGTGAGGTGCTGTGTGAAGGACAACTTCCCCAGCACACAGCAGGAGCAGCA 3807
 QY 1348 -----GlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeuAlaLeuLeu 1364
 DB 3808 ACAACCTCAAAATCTCTCCCTGGCATCTTCTCCCTGGGTGGGTCTCTGCTTATCTCTG 3867
 QY 1365 SerSerIlePheGlyLeuLeuLeuValLeuPheIleLeuPheLeuThrTrpCysArg 1384
 DB 3868 GGTCTGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3924
 QY 1385 ValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArgGlySerLeuGlu 1404
 DB 3925 GCAGAGCGC-----AGAGCCTTATCCAGCTATGAAGATGCTCTGTCT 3966
 QY 1405 GluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspProHisGlyThrArg 1424
 DB 3967 GRAGCTGTGTATGAGAGCTCGATTACCTCTGACACAGAGGAAGGTCTGGCGACGCCA 4026
 QY 1425 ThrSer---AspAspThrProAsnHisGlyCysGluAspAlaSerAspThrSerLeuLeu 1443
 DB 4027 GATCAGATGACTGATCTCCCTGATGAAATATGATGATGATGATGATGATGATGATGATGATG 4086
 QY 1444 GlyValLeuProAsnSerGluAlaThrLys 1453
 DB 4087 GGAACCTCTTCTCCCTCTCAGGGGAATGAG 4116
 RESULT 12
 ANV49652
 ID AAV49652 standard; DNA; 5802 BP.
 XX
 AC AAV49652;
 DT 23-OCT-1998 (first entry)
 XX
 DE Human SRCR protein DNA.

XX Scavenger receptor cysteine rich domain; SRCR; diagnosis; treatment;
 KW nervous system; medullo-blastoma; glioma; breast; detection;
 KW autoantibody; ss.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT CDS 107..5464
 FT /*tag= a
 FT /product= SRCR protein"
 FT /note= "Scavenger receptor cysteine-rich domain
 FT containing protein"
 XX
 PN W09830687-A2.
 XX
 PD 16-JUL-1998.
 XX
 XX 09-JAN-1998; 98WO-DE00096.
 XX
 XX 18-JUL-1997; 97DE-1030997.
 PR 09-JAN-1997; 97DE-1000519.
 XX
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 XX
 XX Mollenhauer J., Poustka A;
 XX
 XX WPI: 1998-399136/34.
 DR P-PSDB; AAW64591.
 XX
 PT Proteins containing scavenger receptor, cysteine rich domain -
 PT useful for diagnosis and treatment of tumours
 XX
 PS Claim 4; Fig 2; 54pp; German.
 XX
 CC This sequence encodes a human protein which contains a SRCR (scavenger
 CC receptor, cysteine-rich) domain. The gene and encoded protein can be used
 CC to diagnose or treat tumours, particularly of the nervous system
 CC (medullo-blastoma or glioma) or breast. The DNA-sequence and probes
 CC derived from it, are used to identify genes that express SRCR-domain
 CC containing proteins, to determine the form in which these proteins exist
 CC and to assess the significance of individual forms on cellular
 CC properties. The protein can be used to detect the presence of
 CC autoantibodies and antibodies which regulate its expression.
 XX
 SQ Sequence 5802 BP; 1253 A; 1614 C; 1576 G; 1359 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 1.37e-146 Length: 5802
 Score: 2435.00 Matches: 533
 Percent Similarity: 46.58% Conservative: 155
 Best Local Similarity: 36.09% Mismatches: 415
 Query Match: 29.92% Indels: 374
 DB: 19 Gaps: 28
 US-09-759-130B-381 (1-1453) x AAV49652 (1-5802)
 QY 43 GlyThrAsp-----LeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr 60
 DB 389 GGATCTGATCTGTTGGCCCTCAGGCTGGTGAATGGAGATGGCAGGTGTGAGGGCCGA 448
 QY 61 ValGluValLysPheGlnGlnTrpGlyThrValCysAspGlyTrpAsnThrThr 80
 DB 449 GTGAGATCTATACCGAGGCTCTGGGGACCGCTGTGTGATGACAGCTGGGACACCAAT 508
 QY 81 AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMet----- 97
 DB 509 GATGCCAACGGTGGTCTGTAGCAGCTGGTGTGGTGGGCCATGTTCAGCTCCAGGAAT 568
 QY 98 PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCys 117
 DB 569 GCCTGGTTGGCCAGGGC-----TCAGGACCCCATTCGCCGTGGATGATGTCGCGTGC 619

QY 118 TyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCys 137
 DB 620 TCAGGACACGAATCTACCTGTGGAGCTGCCCCACAATGGCTGGCTCTCCCAATACTGT 679
 QY 138 TyrHisGlyGluAspValGlyValAsnCysTyr----- 148
 DB 680 EGCCATGGTGAAGATGCTGGTGTATCTCTCAGCTGCCAGCCTCAGTCAACACTCAGG 739
 QY 149 -----GlyGluAlaAsn 152
 DB 740 CCAGAAAGTTGGCCCTGTCTCAGGATATCACCACCTGTACCACAGAAAGATCTGAATCCAGT 799
 QY 153 LeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArgValGluValLysPhe 172
 DB 800 TTGGCCCTCAGGCTGGTGAATGGAGGCGACAGGTTGTGAGGCCCGCAGTGGAGTCTTATAC 859
 QY 173 GlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaVal 192
 DB 860 CGAGGCTCCTGGGCGACCGTGTGTGATGACTTCTGGGACACCAATGATGCCAATGTGTC 919
 QY 193 CysArgGlnLeuGlyCysProSerSerPheIleSerSerGlyValValAsnSerProAla 212
 DB 920 TGCAGGCGAGCTGGGCTGTGGCTGGCCATGTTCAGCCCGAGGAATGCCAGTTGGCCAG 979
 QY 213 ValLeuArgProIleTrpLeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrp 232
 DB 980 GGCTCAGGACCCATGTCTGGATGTGGCTGCTCAGGACACGAGTCTCTACCTGTGG 1039
 QY 233 AsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHisAsnGluAspValThrLeu 252
 DB 1040 AGCTGCCCGCCACAATGGCTGGCTCACCACCAACTGTGGCCATAGTAGAGAGCGTGGTGC 1099
 QY 253 ThrCysTyr----- 255
 DB 1100 ATCTGCTCAGCTCCCGAGTCCCGCGCGACACCCAGCCAGCATACTTGGCCACCTCAGAT 1159
 QY 256 -----AspSerSerAspLeuGluLeuArgLeuValGlyGlyThrAsnArg 270
 DB 1160 GCATCAACAGCAGCAGCACCTGAATCCAGTTTGGCCCTCAGGCTGGTGAATGGAGGTGACAGG 1219
 QY 271 CysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyThrValCysHisLys 290
 DB 1220 TGTCCAGGCGCGAGTGGAGGCTCTATACCGAGGCTCCTGGGCGCAGCTGTGTGATGATAGC 1279
 QY 291 TrpAsnAsnAlaAlaAspValValCysLysGlnLeuGlyCysGlyThrAlaLeuHis 310
 DB 1280 TGGACACACAGTGACGCCAATGTGTCTCGCGCAGCTGGGCTGGCTGGCCACGCTCA 1339
 QY 311 PheAlaGlyLeuProHisLeuGlnSerGlySerAspValValTrpLeuAspGlyValSer 330
 DB 1340 GCCCCAGGAAATGCCCGGTTTGGCCAGGTTTCAGGACCCCAATTCCTGGATGACGTGGCG 1399
 QY 331 CysSerGlyAsnGluSerPheLeuTrpAspCysArgHisSerGlyThrValAsnPheAsp 350
 DB 1400 TGCTCAGGCTATGATGCTCTACCTGTGGAGCTGCCCGCCACAATGGCTGGCTCCCATTAAC 1459
 QY 351 CysLeuHisGlnAsnAspValSerValIleCysSerAspGlyAlaAspLeuGluLeuArg 370
 DB 1460 TGTCCAGCAGCTGAAGACGCTGGTGTCTATCTGCTCAGCTGCC----- 1501
 QY 371 LeuAlaAspGlySerAsnAsnCysSerGlyArgValGluValArgIleHisGluGlnTrp 390
 DB 1502 -----CACTCCCTGG 1510
 QY 391 TrpThrIleCysAspGlnAsnTrpLysAsnGluAlaLeuValValCysLysGlnLeu 410
 DB 1511 TCGACG----- 1516
 QY 411 GlyCysProPheSerValPheGlySerArgAlaLysProSerAsnGluAlaArgAsp 430
 DB 1517 -----CCAGTCCAGCACACATGTCGCG 1537
 QY 431 IleTrpIleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyr 450

Db 1538 ACCATCACTTCGCCGTCAGACAGTAGATCTGAATCCAGT----- 1579
Qy 451 AspGlyLysAlaLysArgThrCysPheArgArgSerAspAlaGlyValIleCysSerAsp 470
Db 1579 ----- 1579
Qy 471 LysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTyrGlyArgLeuGlu 490
Db 1580 ----- 1580
Qy 491 ValLysTyrGlnGlyClnTrpGlyThrValCysHisAspArgTyrSerThrArgAsnAla 510
Db 1631 GTCCATACCAAGGTCCTGGGGCACCCTGGTGCATGACAGCTGGGACACCAATGATGCC 1690
Qy 511 AlaValValCysLysGlnLeuLeuCysGlyLysProMetHisValPheGlyMetThrTyr 530
Db 1691 AATGTCGTCGACGACCCGGCTGGCTGGCCATGTCAGCCCGAGGAATCCCGG 1750
Qy 531 PheLysGluAlaSerGlyProIleTrpLeuAspValSerCysIleCysGlnSer 550
Db 1751 TTTGGTCAGGGCTCAGGACCCATTCCTGGATGATGTGCGCTGCTCAGGACACAGTCT 1810
Qy 551 AsnIleTrpAspCysGlnHisSerGlyTrpGlyLysHisAsnCysValHisArgGluAsp 570
Db 1811 TACCCGTGGAGTGGCCCAATGGCTGGCTCTCCCAACTGTGGCCATAGTAGAAGAC 1870
Qy 571 ValIleValThrCysSerGlyAspAla-----ThrTrp--- 581
Db 1871 GCTGGTGTCTATCTGCTCAGCTCCAGTCCCGGCCCAACACCTAGTCCAGACACTTGGCCA 1930
Qy 582 ----- 582
Db 1931 ACCTCACATGATCAACAGCAGGATCTGAATCCAGTTTGGCCCTGAGGCTGTGAATGA 1990
Qy 589 SerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgTrpGlyThrValCys 608
Db 1991 GTGACAGGTGTACGGGCGAGTGAGGTCTATACCGAGGCTCCTGGGCGCACCGTGTGT 2050
Qy 609 AspAspGlyTrpAsnSerLysAlaAlaValValCysSerGlnLeuAspCysProSer 628
Db 2051 GATGACTACTGGGACCAACCAATGATGCCAATGTGGTTTCAGGACGCTGGCTGTGGCTGG 2110
Qy 629 SerIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyLysIleTrpLeuAsp 648
Db 2111 GCATGTCTACGCCAGCCAGAAATGCCGGTTTGGCCAGGTTTCCAGGACCCATTCCTCGGAT 2170
Qy 649 AspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgAsnSerGlyTrpGly 668
Db 2171 GATGTGCGTCTCAGGACATGAGTCCTATCTGTGAGCTGCCCAACCAATGGCTGGCTC 2230
Qy 669 AsnAsnAspCysSerHisSerGluAspValGlyValIleCysSer----- 683
Db 2231 TCCCAACATGTGGCCATCATCAAGACGCTGTGTCTCATCTGCTCAGCTTCCAGTCCCGAG 2290
Qy 684 ----- 684
Db 2291 CCGACACCCAGCCACAGACTTGGCCAACTCCATGTCATCAACAGCAGGATCTGAATCC 2350
Qy 687 AspMetGluLeuArgLeuValGlyLysSerArgCysAlaGlyLysValGluValAsn 706
Db 2351 AGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGACAGGTGTCAGGGCGGAGTGGAGTCTCTA 2410
Qy 707 ValGlnGlyAlaValIleLysLeuCysAlaAsnGlyTrpGlyMetAsnIleAlaGluVal 726
Db 2411 TACCGAGGCTCTGGGGCACCCTGTGTGATGACTACTGGGACCAACCAATGATCCCAATGTG 2470
Qy 727 ValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGluProHisPheThr 746
Db 2471 GTTTCAGGCAGCTGGGCTGTGGCTGGCCACAGTCCAGGAAATGCCCCGGTTTGGC 2530
Qy 747 GluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGlyGluAlaSerLeu 766
Db ----- 766

Db 2531 CAGGTTTCAGGACCCATTCCTCGATGATGTGGCTGCTCAGGACATGAGTCTATCTG 2590
Qy 767 TrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsnMetGluAlaSer 786
Db 2591 TGGAGCTGCCCCACAATGGCTGGCTCTCCACAACTGTGGCCATCATCAAGACGCTGTG 2650
Qy 787 LeuIleCysSerAlaHisArg---GlnProArgLeuValGlyAlaAspMetProCysSer 805
Db 2651 GTCATCTCTCAGCTTCCAGTCCAGTCCAGCGG----- 2680
Qy 806 GlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSerAspPheSer 825
Db 2681 ----- 2707
Qy 826 LeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIleSerLeuSer 845
Db 2707 ----- 2707
Qy 846 ValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPheGlnCysGlu 865
Db 2708 -----TCTGTGTCATCAACAGCA 2725
Qy 866 GlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAspThrCysIle 885
Db 2726 GGATCTGAATCCACTTTGGCCCTG----- 2749
Qy 886 HisSerArgGluValGlyValValCysSerArgTyrThrAspValArgLeuValAsnGly 905
Db 2750 -----AGACTGTGAATGA 2764
Qy 906 LysSerGlnCysAspGlyClnValGluIleAsnValLeuGlyHisTrpGlySerLeuCys 925
Db 2765 GGTGACAGGTGTGAGGCGGAGTGGAGTCTATACCAAGGCTCCTGGGCGCACCGTGT 2824
Qy 926 AspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSerCysGlyThr 945
Db 2825 GATGACTACTGGGACCAACCAATGATGCCAAGTGGTGTGACGACGCTGGCTGTGGCTGG 2884
Qy 946 AlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArgValTrpGlyHis 965
Db 2885 GCATGTGACGCCCGAGAAATGCCAGTTTGGCCAGGCTCAGGACCCATTCCTCGGAT 2944
Qy 966 ArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThrValLeuGly 985
Db 2945 GATGTGCGCTGCTCAGGACACAGTCTTACCTGTGGAGCTGCCCGCACCAATGGCTGGCTC 3004
Qy 986 AlaProProCysIleHisGlyAsnThrValSerValIleCysThrGlySerLeuThrGln 1005
Db 3005 TCCACAACTGTGGCCATCATCAAGTGTGTGTCTCATCTGCTCAGTCTCAGTCCCAG 3064
Qy 1006 -----ProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaValPro 1023
Db 3065 TCAAGCCCGCAGCCAGATCTTGGCTGACCACTTACCGGCATTCAGCATAGGATCT 3124
Qy 1024 GluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAspSer 1043
Db 3125 GAATCCAGTTTG-----GCTGTGAGGCTGTGAATGAGGTGAC 3163
Qy 1044 ArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAspAsp 1063
Db 3164 AGGTCTCAGGCGGAGTGGAGGTCTCTATCGAGGCTCCTGGGGAACCGTGTGTGATGAC 3223
Qy 1064 GlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAlaPhe 1083
Db 3224 AGCTGGGACACCAATGATGCCAATGTGTCTGACGAGCTGGGCTGGCTGGCCATG 3283
Qy 1084 AsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspLeu 1103
Db 3284 TCGGCCCGCAGGAATGCGCGTTTGGCCAGGCTCAGGACCCATTCCTCGATGATGTG 3343
Qy 1104 AsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHis 1123
Db 3344 CGCTGCTCAGGGAATGAGTCTTACCTGAGCTGCCCGCCCAAGGCTGGCTCAGCCAC 3403

Oy 1124 AspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeu 1143
: : : : :
Db 3404 AACTGTGGCCATCAGAGACGCTGGTCTCATCTGCTCA-----GCCACCAATA 3454
Oy 1144 TyrSerGluThr-----GluThrGlu 1150
: : : : :
Db 3455 AATCTACTACGACAGATGGTGGCATCCAAACAACACTACACCACTGCCAAGACCTCTTCA 3514
Oy 1151 SerCysAlaGlyArgLeuGluValPheTyr---AsnGlyThrTrpGlySer- 1166
: : : : :
Db 3515 AATGTGGTGGCTTCTTA-----TTCTATCCAGTGGGACATCTCCAGCCCATCTTAC 3568
Oy 1166 ----- 1166
Db 3569 CTGCATACCTACCAACAATGCTAAGTGTCTTTGGGAAATAGAAGTGAATCTGGTTAT 3628
Oy 1167 -----ValGlyArgAsnIle----- 1172
Db 3629 CGCATAAACCTGGGCTTCAGTAATCTGAATTTGGAGGACACCACTAACTGCAGTTTGTAT 3688
Oy 1173 -----ThrThrAlaIleAlaGlyIleValCys--- 1181
: : : : :
Db 3689 TATGTTGAATCTTGTATGATCATTTGAATAGCAGTCTCTGCTGGGAAATCTGTAT 3748
Oy 1182 -----ArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeu-----AlaPro 1196
: : : : :
Db 3749 GATACGAGCAATATTTACATCTCTTTACACCAAGATGACCATTTCTCGAAGTGC 3808
Oy 1197 LeuSerLysThrGlySerGlyPheMet---TrpValAspAspIleGlnCysProLysThr 1215
: : : : :
Db 3809 ATCAGTTCCAAAACACTGGCTTTTGGCTTGGTATAC----- 3847
Oy 1216 HisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgIleSerSerProAla 1235
: : : : :
Db 3848 -----TCTTCCCAAGC 3859
Oy 1236 GluGluThrTrpIleThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCys 1255
: : : : :
Db 3860 GATGCCACCTTG-----AGGTTGCTCAATTTAAATTCATCTATGCTATGT 3907
Oy 1256 SerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrp 1275
Db 3908 GCCGGCGGTGAGAAATTTACATGGTGGCACCTGGGGAGAGTTGTGATGACTCCG 3967
Oy 1276 AspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAla 1295
: : : : :
Db 3968 ACCATTCAGGAAGTGGTGGTCTGCAGACAGCTAGGCTGGACGTGCAGTTTCAGCC 4027
Oy 1296 LeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCys 1315
: : : : :
Db 4028 CTTGGAATGTCATATTTTGGCTCTGGCTCTGGCTCCATCACCCTGGACGATGACAGTGC 4087
Oy 1316 LysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysProTrpGlyLysSerAspCys 1335
: : : : :
Db 4088 TCAGGAGGAGGAATCCTCTCTGGCAGTGGCGGAGACCGAGCTGGTCTCCACACATGT 4147
Oy 1336 GlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSerLeuLysSer 1352
: : : : :
Db 4148 AATCATCTGAAGTCTGGTGTCTATCTGCTCAGGAACCATCTATCTGACA 4198
RESULT 13
AAS17590
ID AAS17590 standard; cDNA: 2697 BP.
XX
AC AAS17590;
XX
DT 26-FEB-2002 (first entry)
XX
DE DNA encoding novel secreted protein #19.
XX
KW Secreted protein; cytostatic; immunosuppressive; vulnery; vaccine;
antlinflammatory; neuroprotective; nephrotropic; cardiovascular;

KW human; cancer; autoimmune disease; wound healing disorder; infection;
KW haematopoietic disorder; inflammatory disorder; infertility;
KW neurological disease; psychiatric disease; cardiovascular disease;
KW respiratory disease; renal; gastrointestinal; ss.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 1..2697
FT /*tag= a
FT /product= "Human secreted protein"
XX WO200179454-A1.
PN
XX 25-OCT-2001.
PD
XX 11-APR-2001; 2001WO-US11797.
PF
XX 13-APR-2000; 2000US-196603P.
PR 24-APR-2000; 2000US-199417P.
PR
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX (SMIK) SMITHKLINE BEECHAM PLC.
PA
XX Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z;
PI
XX WPI: 2002-061975/08.
XX P-PSDB: RAU09878.
DR
XX New secreted proteins or polypeptides, useful for treating e.g. cancer,
PT autoimmune diseases, wound healing disorder, infections, haematopoietic
PT disorders, inflammatory disorders, infertility, cancer.
XX
PS Claim 2; Page 47-48; 92pp; English.
XX

CC The invention relates to an isolated novel secreted polypeptide (I) and
CC polynucleotide (II). (I) and (II) are useful for treating cancer,
CC autoimmune diseases, wound healing disorder, infections, haematopoietic
CC disorders, inflammatory disorders, infertility, neurological and
CC psychiatric diseases, cardiovascular diseases, respiratory diseases,
CC renal diseases, or gastrointestinal diseases. These may also be used to
CC treat diseases, abnormalities and disorders caused by abnormal
CC expression, production, function and/or metabolism of the genes, as
CC vaccines for inducing immunological response in a mammal, and in
CC screening methods for detecting the effect of added compounds on the
CC secretion of mRNA and polypeptide in cells. The polypeptides can be used
CC as immunogens to produce antibodies immunospecific for the polypeptides,
CC and to identify membrane-bound or soluble receptors. The polynucleotides
CC may be used as diagnostic reagents, in chromosome localisation studies,
CC and in tissue expression studies. The present sequence represents the
CC coding sequence of novel human secreted protein #19.
XX

SQ Sequence 2697 BP; 358 A; 841 C; 1055 G; 443 T; 0 other;

Alignment Scores:
Pred. NO.: 6.94e-104 Length: 2697
Score: 1763.50 Matches: 361
Percent Similarity: 49.37% Conservative: 108
Best Local Similarity: 38.00% Mismatches: 262
Query Match: 21.67% Indels: 219
DB: 24 Gaps: 16

US-09-759-130B-381 (1-1453) x AAS17590 (1-2697)

Oy 536 GlyProIleTrpLeuAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCys 555
Db 31 GGGCCCTCTCTCTG-----AATCTCTGGGAGTC 60
Oy 556 GluHisSerGlyTrpGlyLysHisAsnCysValHisArgGluAspValIleValThrCys 575
Db 61 CCCATTGGTGACACAGT----- 78
Oy 576 SerGlyAspAlaThrTrpGlyLeuArgLeuValGlySerAsnArgCysSerGlyArg 595

Db 386 TGAGTCAGCTCTCTGGGAATGTCACACCGGGAATGGGAAGCCATACTGTTATCATGG 445
 Qy 140 yGluAspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspG1 160
 Db 446 AGAAGAGTGTGTGAAGTGTAACTGTTACGGTGAAGCCATCTGGGTTTGAGGCTAGTGATGG 505
 Qy 160 yAsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTyrGlyThrIleCy 180
 Db 506 AAACAC-TCCTGTTTCAGGAGAGTGGAGTGAATTCACAAAGAGTGGGAACATATATG 564
 Qy 180 sAspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSe 200
 Db 565 TGATGATGGTGGAACTTAAATACCCCTCCGCTGTCGAGGCAACTAGGATGCCATC 624
 Qy 200 rSerPheIleSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAs 220
 Db 625 TCTCTTTATCTCTGAGTGTCTAACAGCCCTGCTGTATGCGCCCATTTGGCTGGA 584
 Qy 220 pAspIleLeuGlnGlnGlyAsnGlnLeuAlaLeuTrpAsnCysArgHisArgGlyTrpG1 240
 Db 685 TGACATTTATGCCAGGGAATGAGTGGCACTCTGGAATTCGACACATCTGGATGGG 744
 Qy 240 yAsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLe 260
 Db 745 AAATCATGACTCAGTCACATGAGGATGTCACATTAATCTGTTATGATAGTAGTCT 804
 Qy 260 uGluLeuArgLeu 264
 Db 805 TGAACGTAGGCTC 817

RESULT 15

AAH99258
 ID AAH99258 standard; cDNA; 690 BP.

AC AAH99258;

DT 16-OCT-2001 (first entry)

DE Human protein encoding cDNA sequence SEQ ID NO:93.

KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
 KW antinflammatory; antirheumatic; antiarthritic; immunosuppressive;
 KW antibacterial; endocrine; cardiant; central nervous system; virucide;
 KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
 KW antiagregant; haemostatic; vulnery; antiulcer; osteopathic; eczema;
 KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
 KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
 KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
 KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmune;
 KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
 KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KW neurological disorder; ss.

OS Homo sapiens.

PN WO200153455-A2.

XX 26-JUL-2001.

PF 22-DEC-2000; 2000WO-US35017.

XX 23-DEC-1999; 99US-0471275.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Drmanac RT;

PI

XX

DR WPI; 2001-457603/49.
 XX P-PSDB; AAM25317.
 PT Isolated human polynucleotides encoding polypeptides, useful for the
 XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
 PS Claim 1; Page 341; 1217pp; English.
 XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
 CC AAM25963. The proteins can have activities based on the tissues and
 CC cells they are expressed in, such as: antinflammatory; antirheumatic;
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
 CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antianaemic; antiagregant; haemostatic; vulnery;
 CC antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.
 XX Sequence 690 BP; 163 A; 134 C; 207 G; 186 T; 0 other;

Alignment Scores:

Pred. No.: 2,42e-67 Length: 690
 Score: 1183.00 Matches: 214
 Percent Similarity: 99.07% Conservative: 0
 Best Local Similarity: 99.07% Mismatches: 2
 Query Match: 14.54% Indels: 0
 DB: 22 Gaps: 0

US-09-759-130B-381 (1-1453) x AAH99258 (1-690)

Qy 636 GlyAsnAlaSerThrGlyTyGlyLysIleTrpLeuAspValSerCysAspGlyAsp 655
 Db 43 GGAACGGCTCTACAGGATATGGAATAATTTGGCTCGATGATGTTCTTCGTGATGGAGAT 102
 Qy 656 GluSerAspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSer 675
 Db 103 GAGTCAGATCTCTGCTCATGCAGGAACAGTGGTGGGAAATAATCACTGCAGTCACACT 162
 Qy 676 GluAspValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGly 695
 Db 163 GAAGATGTTGGAGTATCTGTTCTGATGCATCGGATATGGAGCTTGGGCTTGGGTGGA 222
 Qy 696 SerSerArgCysAlaGlyLysValGluValAsnValGlnGlnAlaValGlyIleLeuCys 715
 Db 223 ACCACAGGTGTGCTGGGAAGTGAACATTGCTGAAGTTGTTTTCAGGCAACTTCAATGTGGTCT 282
 Qy 716 AlaAsnGlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySer 735
 Db 283 GCTAATGGCTGGGAATGAACATTGCTGAAGTTGTTTTCAGGCAACTTCAATGTGGTCT 342
 Qy 736 AlaIleArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSer 755
 Db 343 GCAATCAGGGTCTCCAGAGAGCCTCATTTACAGAGAAGAACATTACACATCTAATGTCC 402
 Qy 756 AsnSerGlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLys 775
 Db 403 AATTCTGGCTGCTGGAGGGAAGCCTCTCTCTGGGATGTATACGATGGAGTGGGAAA 462
 Qy 776 GlnThrAlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnPro 795
 Db 463 CAGACTGCGTGTCAITTAATATATGAAGCAAGTTTGATCTGCTCAGCCACAGCCAGCCC 522

Qy 796 ArgLeuValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThr 815
|||||
Db 523 AGGCTGGTTGGAGCTGATATGCCCTGCTCTGGACGTGTGAAGTGAACATGCACACACA 582
|||||
Qy 816 TrpArgSerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGlu 835
|||||
Db 583 TGGCGCTCTGTCTGTGATTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAA 642
|||||
Qy 836 LeuAsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGly 851
|||||
Db 643 TTAACCTGTGGAGATGCCATATCTTCTCTGGGAGATCATTGGG 690
|||||

Search completed: May 12, 2003, 06:58:24
Job time : 670 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 12, 2003, 10:35:25 ; Search time 2862 Seconds
(without alignments)
8222.237 Million cell updates/sec

Title: US-09-759-130B-381
Perfect score: 8138
Sequence: 1 MMLPONSWHIDFGRCCHON.....CEDASDTSLGLVPASEATK 1453

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus.p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09759130/runat_06052003_123554_23365/app_query.fasta_1.1607
-DB=EST -QFW=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=BIOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09759130.ecgn_1.1.2168 @runat_06052003_123554_23365 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estlin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_esttc.*
9: gb_estti.*
10: gb_est2.*
11: gb_est3.*
12: gb_est4.*
13: gb_est5.*
14: gb_est6.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_oth.*
26: em_gss_pro.*
27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1765.5	21.7	999	9	AL529439	AL529439 AL529439
2	1083	13.3	940	9	AL550994	AL550994 AL550994
3	1072	13.2	582	12	BG570972	BG570972 602591546
4	960	11.8	799	13	BI185454	BI185454 UNL-P-FN-
5	942.5	11.6	1108	9	AL544534	AL544534 AL544534
6	936.5	11.5	1100	9	AL540905	AL540905 AL540905
7	929	11.4	718	13	BI183796	BI183796 UNL-P-FN-
8	895	11.0	554	9	AL529438	AL529438 AL529438
9	836.5	10.3	930	13	BI838644	BI838644 603086185
10	767	9.4	1153	9	AL540904	AL540904 AL540904
11	739	9.1	723	13	BI836459	BI836459 603082833
12	718	8.8	644	13	BI009838	BI009838 MR4-RT004
13	701	8.6	585	14	BQ327934	BQ327934 MR4-RT004
14	695	8.5	653	10	BE502724	BE502724 hz78412.x
15	679	8.3	616	13	BI009795	BI009795 MR4-RT004
16	675	8.3	652	10	BB622216	BB622216 BB622216
17	669.5	8.2	913	14	BQ710454	BQ710454 AGENCOURT
18	667.5	8.2	711	14	BQ378694	BQ378694 MR3-UT009
19	667	8.2	574	14	BQ327932	BQ327932 MR4-RT004
20	666	8.2	581	14	BQ776070	BQ776070 UT-H-FHO-
21	660.5	8.1	593	14	BQ327876	BQ327876 MR4-RT004
22	658	8.1	574	13	BI009286	BI009286 MR4-RT004
23	657	8.1	583	14	BQ327909	BQ327909 MR4-RT004
24	651	8.0	596	13	BI009802	BI009802 MR4-RT004
25	649	8.0	934	14	BQ936228	BQ936228 AGENCOURT
26	645.5	7.9	927	9	AL570629	AL570629 AL570629
27	634.5	7.8	689	12	BF096204	BF096204 MR3-UT009
28	634	7.8	2925	11	AK014331	AK014331 Mus muscu
29	626	7.7	505	13	BI038644	BI038644 IL3-NT028
30	626	7.7	705	17	BI6308	BI6308 34707.TP.CI
31	625	7.7	607	14	BQ327933	BQ327933 MR4-RT004
32	624	7.7	3665	11	BC035470	BC035470 Homo sapi
33	621	7.6	523	12	BG334729	BG334729 602533892
34	620	7.6	569	10	AW631841	AW631841 90867.MAR
35	616.5	7.6	833	9	AL574811	AL574811 AL574811
36	616	7.6	560	13	BI009799	BI009799 MR4-RT004
37	612	7.5	529	13	BI009105	BI009105 MR4-RT004
38	612	7.5	531	14	BQ327873	BQ327873 MR4-RT004
39	612	7.5	579	14	BQ327906	BQ327906 MR4-RT004
40	607	7.5	602	12	BG641637	BG641637 PG11c.pko
41	605	7.4	583	14	BQ368412	BQ368412 PM3-GN051
42	587.5	7.2	594	10	AV616074	AV616074 AV616074
43	579	7.1	822	12	BF792807	BF792807 602253274
44	578	7.1	1581	11	BC020481	BC020481 Homo sapi
45	577.5	7.1	850	9	AL550308	AL550308 AL550308

ALIGNMENTS

RESULT 1
AL529439
LOCUS
DEFINITION AL529439 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD006Y005 5
prime, mRNA sequence.
ACCESSION AL529439
VERSION AL529439.1 GI:12792932
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 999)
AUTHORS Li,W.B., Gruber,C., Jesse,J., and Polayes,D.
TITLE Full-length cDNA libraries and normalization

JOURNAL
COMMENT

Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. .999

FEATURES
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDD006YC05"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 220 a 236 c 318 g 224 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 5,88e-168 Length: 999
Score: 1765.50 Matches: 327
Percent Similarity: 98.50% Conservatives: 1
Best Local Similarity: 98.20% Mismatches: 5
Query Match: 21.69% Indels: 3
DB: 9 Gaps: 0

US-09-759-130B-381 (1-1453) x AL529439 (1-999)

QY 1060 ileCysAspArgpTrpAspLeuSerAspAlaHisValValcysGlnLysLeuGlyCys 1079
Db 1 ATCTGTGATGACGGCTGGGACCTGACCGATGCCAGTGGTGTGTCACAAAGCTGGGCTGT 60
QY 1080 GlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerClyProIleTrp 1099
Db 61 GGAGTGGCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTGAGGCCCATCTGG 120
QY 1100 LeuAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGly 1119
Db 121 CTGGATGACCTGAACCTGCACAGGAATGGAGTCCCACTTGTGGCAGTGCCTTCCCGCGGC 180
QY 1120 TrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThr 1139
Db 181 TGGGGCAGCAGCACTGCAGGCACAGGAGGACGCGAGGGGTCTATCTGCTCAGAATTCACA 240
QY 1140 AlaLeuArgLeuTyTrSerGluThrGluThrGluSerCysAlaGlyValGluValPhe 1159
Db 241 GCCTTGAGGCTCTACAGTGAACCTGAACACAGAGAGCTGTGCGGAGATGGAGTCTTC 300
QY 1160 TyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrAlaIleAlaGlyIle 1179
Db 301 TATAACGGACCTGGGGCAGCGTGGCAGGAGGAACATCACACAGCCATAGCAGGCATT 360
QY 1180 ValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLys 1199
Db 361 GTGTGAGCAGCAGTGGGCTGTGGGGAATGGAGTAGTCAGCTCGCCCTTATCTAAG 420
QY 1200 ThrGlySerClyPheMetTrpValAspIleGlnCysProLysThrHisIleSerIle 1219
Db 421 ACAGGCTCTGGTTTCAATGTTGGTGGATGACATTCAGTGTCTTAAACGCAATATCTCATA 480
QY 1220 TrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrp 1239
Db 481 TGGCAGTGGCTGTCTGCCCCATGGGACGAGAAATCTCCAGCCCGCAGCAGAGACCTGG 540

QY 1240 IleThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgVal 1259
Db 541 ATCACATGTGAAGATAGATAAAGAGTGGTGGAGGAGACACCGAGTCTCTGGGAGAGTG 600
QY 1260 GluIleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGlu 1279
Db 601 GAGATCTGGCAGCGCAGGCTCTCTGGGGCACAGTGTGTGATGACTCTGGGACCTGGCCGAG 660
QY 1280 AlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAla 1299
Db 661 GCGGAAGTGGTGTGTACAGCAGCTGGGCTGTGGCTCTGCTGCTGGCTGCCCTGAGGACGCT 720
QY 1300 SerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGlu 1319
Db 721 TCGTTTGGCCAGGGAACCTGGGAACCATCTGGTTGGATGACATCGCTGGTGGCAAGAAATGAG 780
QY 1320 SerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGlu 1339
Db 781 TCATTTCATGGAGCTGTACGCCAACCCTGGGACAGAGTGTGGACACCAAGNA 840
QY 1340 AspAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHis 1359
Db 841 GATGCTGCGGTGAGTGTCTGGACAGTGGTGAATCACTGAATGCCCTCCTCAGGTGAT 900
QY 1360 LeuAlaLeuLeuSerSerIlePheGlyLeuLeuLeuValLeuPheIleLeuPhe 1379
Db 901 TTAGCACTTATTTTATCCAGTATCTTGGGCT-CTTCTCTGGTCTGTATAT-CTATTT 958
QY 1380 LeuThrTrpCysArgValGlnLysGlnLysHisLeuPro 1392
Db 959 CTCACGTGGTGGCGAGTTCAGAA-WAAAAACATCTGCC 996
RESULT 2
AL550994 940 bp mRNA linear EST 16-FEB-2001
LOCUS AL550994 LTI_NFL006_PL2 Homo sapiens cDNA clone CSODI066YE07 5
DEFINITION AL550994 prime, mRNA sequence.
ACCESSION AL550994
VERSION AL550994.1 GI:12888516
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 940)
AUTHORS Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. .940
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODI066YE07"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site:1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 228 a 212 c 279 g 218 t 3 others
ORIGIN

Alignment Scores:

Pred. No.: 1,03e-98 Length: 940
 Score: 1083.00 Matches: 190
 Percent Similarity: 75.24% Conservative: 41
 Best Local Similarity: 61.89% Mismatches: 76
 Query Match: 13.31% Indels: 1
 DB: 9 Gaps: 0

US-09-759-130B-381 (1-1453) x AL550994 (1-940)

Qy 858 TrpAlaGluTyrPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleVal 877
 Db 9 TGGGCTGAAGATTCAGTGTGAGGACATGAGTCCCATCTTTCACCTCTGCYCATGACA 68
 Qy 878 GlnHisProGluAspThrCysIleHisSerArgValGlyValValCysSerArgTyr 897
 Db 69 CCCCGCCAGAGAACTGTGAGCCAGACAGGAGGATGTGGAGTAGTCTGCTCAAGATAC 128
 Qy 898 ThrAspValArgLeuValAsnGlyLysSerClnCysAspGlnValGluIleAsnVal 917
 Db 129 ACAGAAATCCCTTGGTGAATGGCAAGACCCCGTGTGAGGACAGAGTGGAGCTCAAAACG 188
 Qy 918 LeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeu 937
 Db 189 CTGGTGCCTGGGATCCCTCTGTAACCTCCTCAGTGGACATAGAGATGCCCATGTCTT 248
 Qy 938 CysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGlu 957
 Db 249 TGCCAGCAGCTAAATGTGGAGTTGCCCTTCTACCCAGGAGGACGACGTTTGGAAA 308
 Qy 958 ArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGlnSerLeuLeuAsp 977
 Db 309 GGAATGGTCAGATCTGGAGCATATGTTTCACCTGCATGGAGTGGAGCAGCAGCATGGGA 368
 Qy 978 AsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerVal 997
 Db 369 GATGTCTGTACTGTCTAGTGTCTCATATGCTTTCAGACAGAGTGGCCCTCTGTA 428
 Qy 998 IleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValIserAspPro 1017
 Db 429 ATCTGCTCAGGAAACCACTGCCAACACTGTCCTCGTGCATTCATGCTCTTGGGCCCA 488
 Qy 1018 TyrLeuSerAlaValProGluGlySerAlaLeuLeuLeuLeuGlyAspLysArgLeuArg 1037
 Db 489 ACAAGGCTACCATCCAGAAAGTGTGTGTCCTGCATAGAGAGTGGTCAACTTCGC 548
 Qy 1038 LeuValAspGlySerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrp 1057
 Db 549 CTGTAATGGAGGAGTGGCTGTCTGGGAGTAGAGATCTATCATGAGGCTCTCTGG 608
 Qy 1058 GlyThrIleCysAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeu 1077
 Db 609 GGCACCATCTGTATGACAGCTGGGACCTGAGTGCATGCCACGCTGTTGCAGACAGCTG 668
 Qy 1078 GlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyPro 1097
 Db 669 GGCTGTGGAGAGGACATTAATGCCACTGGTCTCTCTATTTGGGGAAGAAACAGGCGCC 728
 Qy 1098 IleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSer 1117
 Db 729 ATCTGGCTGGATGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 788
 Qy 1118 ArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGlu 1137
 Db 789 CACGCTGGGGGACGAAATGTCAGGCACACAGGAGGATGGGAGATATCTGCTCAGAA 848
 Qy 1138 PheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGlu 1157
 Db 849 TTCATGCTCTCTKAGACTGACCAAGTGAAGCCAGCAGAGAGGCGCTGTGACGGSGTCTGAA 908
 Qy 1158 ValPheTyrAsnGlyThrTrp 1164
 Db 909 GT-TTTTACATGAGCTTGG 928

RESULT 3
 BG570972

LOCUS 602591546F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4714115 5',
 DEFINITION mRNA sequence.

ACCESSION BG570972 GI:13578625
 VERSION BG570972.1
 KEYWORDS EST.

SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 582)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaaps@email.nih.gov

Tissue Procurement: CLONTECH Laboratories, Inc.
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: L1CMI558 row: k column: 12
 High quality sequence stop: 576.

FEATURES
 source 1..582
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4714115"
 /lab_host="NIH_MGC_79"
 /note="Organ: placenta; Vector: pDNR-LIB (Clontech);
 Site_1: SfiI (ggccgctcgcc); Site_2: SfiI (ggccattatggcc
); 5' and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor
 sequence: 5'-ATTCTAGAGCGGAGCGGCGGACATG-dT(30)BN-3',
 (where B = A, C, or G and N = A, C, G, or T). Average
 insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
 Library."

BASE COUNT 126 a 140 c 198 g 118 t
 ORIGIN

Alignment Scores:
 Pred. No.: 5,38e-98 Length: 582
 Score: 1072.00 Matches: 192
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.48% Mismatches: 0
 Query Match: 13.17% Indels: 0
 DB: 12 Gaps: 0

US-09-759-130B-381 (1-1453) x BG570972 (1-582)

Qy 1109 GluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLys 1128

Db 2 GAGTCCCACTTGTGGCAGTGCCTTCCCGCGCTGGGGCAGCAGCAGCTGCAGGCACAAG 61

Qy 1129 GluAspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGlu 1148

Db 62 GAGACCGAGGGGTCTCTGCTCAGAAATTCACGCTTGAGGCTCTACAGTGAACCTGAA 121

Qy 1149 ThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGly 1168

Db 122 ACAGAGAGCTGTCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCGTGGC 181

Qy 1169 ArgArgAsnIleThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGlu 1188

Db 182 AGGAGGAACATCACACAGGCATAGCAGGCAATTGTGTGACGAGCAGCTGGGCTGTGGGAG 241

LOCUS AL544534 1108 bp mRNA linear EST 16-FEB-2001
 DEFINITION AL544534 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1021YH23 5
 prime, mRNA sequence.
 ACCESSION AL544534
 VERSION AL544534.1 GI:12877014
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1108)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 FEATURES
 Location/Qualifiers
 1. .1108
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0D1021YH23"
 /clone_lib="LTI_NFL006_PL2"
 /tissue_type="placenta"
 /notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact : Feng Liang Life Technologies,
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"
 BASE COUNT 307 a 189 c 319 g 289 t 4 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 2 58e-84 Length: 1108
 Score: 942.50 Matches: 169
 Percent Similarity: 65.58% Conservative: 52
 Best Local Similarity: 50.15% Mismatches: 111
 Query Match: 11.58% Indels: 5
 DB: 9 Gaps: 3
 US-09-759-130B-381 (1-1453) x AL544534 (1-1108)
 Qy 1 MetMetLeuProGlnAsnSerTrpHisIleAspPheGlyArgCysCysHis---Gln 19
 Db 69 ATGGTGCTACTTGAAGACTCTGGATCTGCTGACTTCAGAGA-----CATTTTGTG 119
 Qy 20 AsnLeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSer 39
 Db 120 AACTTGAGTCCCTTCACATTACTGTGCTTACTTCTAGTGGCTGTGTGTGTCACCACT 179
 Qy 40 SerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGly 59
 Db 180 TCTCTTGGAGGACAGACAGAGGACTGAGGCTAGTGGATGGTGAACAACTGTAGCGNG 239
 Qy 60 ThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThr 79
 Db 240 AGAGTGAAGTGAAGTCCAGGAGGAGTGGGAACTGTCTAATAATGCTGGAGCATG 299
 Qy 80 ThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArg 99
 Db 300 GAACCGGTCTCTGTGATTTGTAACCACTGGGATGTCCAACTGATCAAGCCCTCGGA 359
 Qy 100 PheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspValSerCysTrpGly 119
 Db 360 TGGCTAATTCAGTGGAGGTTCTGGACGCAATTTGGATGATCATGTTCTTGTCTGGGG 419
 Qy 120 AsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHis---AsnCysTrp 138

Db 420 AATGAGTCAGCTCTTTGGGATTGCAACATGATGGATGGGAAAGCATAGTAAGTACT 479
 Qy 139 HisGlyGluAspValGlyValAsnCysTrpGlyGluAlaAsnLeuGlyLeuArgVal 158
 Db 480 CACCAACAAGATGCTGGAGTGACCTGCTCAGATGGATCCCAATTTGGAAATGAGGCTGACG 539
 Qy 159 AspGlyAsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThr 178
 Db 540 CGTGGAGGGAATATGCTTCTGGGAAGATAGATCAAAATTCGAAGACGGTGGGGAACA 599
 Qy 179 IleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCys 198
 Db 600 GTGTGTGATGATAAATTCACATAGATCATGCTCTGTCTGTAGCAAACTTGAATGT 659
 Qy 199 ProSerSerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrp 218
 Db 660 GGAAGTCTGTGCTGCTTCTCTGTTTCATTAATTTGGAGAGGCTCTGGACCAATCTGG 719
 Qy 219 LeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGly 238
 Db 720 TTTGATGATCTTATATGCAAGGGAATGAGTCAGCTCTCTGGAATGCAAAATCAAGGA 779
 Qy 239 TrpGlyAsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTrpAspSerSer 258
 Db 780 TGGGAAAGCATAAATCTGTGATCATGCTCAGGATGCTGAGTGTGCTCAAGGAGCA 839
 Qy 259 AspLeuGluLeuArgLeuValGlyThrAsnArgCysMetGlyArgValGluLeuLys 278
 Db 840 GATCTGACCTGACGAGTGGTAGATGGAGTCACTGAATGTTCCAGGAAGATTAGAAGTGA 899
 Qy 279 IleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAspVal 298
 Db 900 TTCAAGGAGATGGGGACAAATATGTCACCGCTGGGACACTAGGATGCTGCTGTG 959
 Qy 299 ValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGln 318
 Db 960 GATGCAAGCAACTGGGATGTCCAACTCCCTCAGCAGCCATTTGTCGAGTTAAGGCCAT 1019
 Qy 319 SerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGlu 335
 Db 1020 VAGGGATTTGGACACATCTGGCTTGACAGCGCTTTCTTCCAGGACACATGAA 1070
 RESULT 6
 AL540905 1100 bp mRNA linear EST 16-FEB-2001
 LOCUS AL540905 LTI_FL002_PL1 Homo sapiens cDNA clone CS0DE004YI19 5 prime
 DEFINITION , mRNA sequence.
 ACCESSION AL540905
 VERSION AL540905.1 GI:12871463
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1100)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 FEATURES
 Location/Qualifiers
 1. .1100
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DE004YI19"
 /clone_lib="LTI_FL002_PL1"
 /lab_host="DH10B"
 /note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
 cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-stranded cDNA was digested with Not I

ORIGIN

US-09-759-130B-381 (1-1453) x AL540905 (1-1100)

Alignment Scores:

```
Pred. No.: 2,65e-83 Length: 718
Score: 929.00 Matches: 156
Percent Similarity: 82.13% Conservative: 37
Best Local Similarity: 66.38% Mismatches: 42
Query Match: 11.42% Indels: 1
DB: 13 Gaps: 0

US-09-759-130B-381 (1-1453) x B1183796 (1-718)
Qy 1050 Glut1eYrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAsp 1069
Db 14 GAGTCTATCATAGAGGCTCTCGGGGACCACATCTGTGATGACACTGGGACCTGAATGAT 73
Qy 1070 AlaHisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAla 1089
Db 74 GCCATCTGGTGTCAACACAGCTGAGCTGTGGATGGCCATTATGCCACTGGTCTGCT 133
Qy 1090 HisPheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGlu 1109
Db 134 CATTTTGGGAAGGAACAGGCGCCATTTGGCTGGATGAGATAAACTGTAATGGAAAGAA 193
Qy 1110 SerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysAtqHisLysGlu 1129
Db 194 TCTCATATTGGCAATGCCACTACATGTTGGGGGGCGGCAATTCAGGCGCATGAAGAG 253
Qy 1130 AspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGluThr 1149
Db 254 GATCAGGAGTCATCTGCTGGAGTTTCATGCTCTCAGACTGATCAGTGAACAGCAGA 313
Qy 1150 GluSerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArg 1169
Db 314 GAGACCTGTGCGAGGCGCTGGNAGTTTTTTACAAACGAGCTTGGGGCAGCGTTGGCAAG 373
Qy 1170 ArgAsnIleThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsn 1189
Db 374 AATGACATGCTCCAGCCACAGTGGGGTGTATGCGAGCAGCTGGGCTGTGCAGACAGA 433
Qy 1190 GlyValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAsp 1209
Db 434 GGGGACATCAGCCCTGTCATCTTCAGACAAGACAGTGTCCAGGCACATGTGGGTGGACAAT 493
Qy 1210 IleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArg 1229
Db 494 GTTCAGTGTCTTAAAGGACCTTGACACCCATATGGCAGTGCCTCATCTCCATGGAAGAAG 553
Qy 1230 ArgIleSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArgValArg 1249
Db 554 AGACTGGCCACCCCTCAGAGGACATGATGATCATGTGCCAACAAATAAGACTTCAA 613
Qy 1250 GlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThr 1269
Db 614 GAAGNACACATAATTGTTCTNGACGTGTGGAGATCTGGTACGGNGGTTCTCTGGGCACT 673
Qy 1270 ValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCys 1284
Db 674 GTGTGTACGACTCTCTGGGACTTGAA-AATGCNCAAGTGTGTGC 717

RESULT 8
AL529438/c AL529438 554 bp mRNA linear EST 13-FEB-2001
LOCUS AL529438 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD006Y005 3
DEFINITION prime, mRNA sequence.
ACCESSION AL529438
VERSION AL529438.1 GI:12792931
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 554)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
```

COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1. 554
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 132 a 166 c 127 g 110 t 19 others
ORIGIN

Alignment Scores:
Pred. No.: 4,57e-80 Length: 554
Score: 895.00 Matches: 164
Percent Similarity: 90.81% Conservative: 4
Best Local Similarity: 88.65% Mismatches: 16
Query Match: 11.00% Indels: 1
DB: 9 Gaps: 0

US-09-759-130B-381 (1-1453) x AL529438 (1-554)

Qy 1189 AsnGlyValValSerLeuAlaProLeu-SerLysThrGlySerGlyPheMetTrpValAs 1208
Db 554 AATGGAGTTATCAGCTCGCCCTTTATTTCTAAGACAGGCTCTGGTTTCATGTGGGTGGA 495
Qy 1208 PASIleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpG 1228
Db 494 TRACATTCACTGTCTTAAACGATGCTCTCCATATGGCAGTGCCTTCTGCCCATGGGA 435
Qy 1228 uArgArgIleSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArgVa 1248
Db 434 CGGAAGACTCTCCAGCCCGVCGAGAGACCTGGCTMACATGTGAAGATAGATAGAGT 375
Qy 1248 lArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpG 1268
Db 374 CGGTGGAGGACACCGAGTCTCTGGGAGAGTGGAGATCTGGCAGCAGGCTCTCTGGGG 315
Qy 1268 yThrValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuG 1288
Db 314 CACAGTGTGTATVACTCTCTGGGACCTGGCGGAGCGAGNAGTGTGTNTCAGCAGCTGG 255
Qy 1288 yCysGlySerAlaLeuAlaLeuAlaLeuAtqAspAlaSerPheGlyGlnGlyThrGlyThr 1308
Db 254 CTGTGGCTCTGCTCTNGCTGCTCCCTGAGGAGCAGCTCTGTTTGGCCAGGGGACTGGAC 195
Qy 1308 eTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaL 1328
Db 194 CTGGTTGGNTTACATGCGGTGCAAAAGGAANTNAGTCATTTCTATGGGACTGTGCACG 135
Qy 1328 sProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGly 1348
Db 134 ACCCTGGGAGACAGTACTGTGGACACAAGGAGANGCTGGCGTGTAGTGTCTCTGGACA 75
Qy 1348 nSerLeuLysSerLeuAsnAlaSerSerGlyHisLeuAlaLeuLeuLeuSerIlePhe 1368
Db 74 CTCGCTGNAATCAATGAATGCTCTCTCAGGCTATTTAGGACTTATTTTATCAGTGTCT 15
Qy 1368 eGlyLeuLeuLeu 1372

/db_xref="taxon:9606"
 /clone="CS0DE004Y119"
 /clone_lib="LTI_FL002_PL1"
 /lab_host="DH10B"
 /note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
 cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-stranded cDNA was digested with Not I
 and cloned into the Not I and Eco RV sites of the
 pCMVSPORT 6 vector. Library was constructed by Life
 Technologies. Contact: Feng Liang Life Technologies, a
 division of Invitrogen 9800 Medical Center Drive Rockville
 , Maryland 20850, USA Fax: (1) 301 610 8371 Email:
 fliang@lifetech.com URL:
 http://fulllength.invitrogen.com"

BASE COUNT 292 a 300 c 228 g 313 t 20 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1.83e-66 Length: 1153
 Score: 767.00 Matches: 146
 Percent Similarity: 65.45% Conservative: 51
 Best Local Similarity: 48.50% Mismatches: 88
 Query Match: 9.42% Indels: 16
 DB: 9 Gaps: 5

US-09-759-130B-381 (1-1453) x AL540904 (1-1153)

Qy 1164 TrpGlySerValGlyArgArgAsnIleThrAlaIleAlaGlyIleValCysArgGln 1183
 Db 1153 TGGGGCACTGTGGCAAGAGCAATGCTGAAMCCACTGTGGGTGGTGGCAGGCG 1094
 Qy 1184 LeuGlyCysGlyGluAsnGlyValSerLeuAlaProLeuSerLysThrGlySerGly 1203
 Db 1093 CTGGCTGTGCAGCAAGGGAATATCACCCTGCTTGTAGACAAGCCATGTCATT 1034
 Qy 1204 PheMetTrpValAspIleCysProLysThrHisIleSerIleTrpGlnCysLeu 1223
 Db 1033 CCCATGTGGGTGGCAATGCTCAGTGTCCAAAGGACCTGCACACGTTGGCAGTCCCA 974
 Qy 1224 SerAlaProTrpGluArgIleSerSerProAlaGluGluThrTrpIleThrCysGlu 1243
 Db 973 TCATCTCCATGGGAAGAGAGTGGCCAGCCCTCGGAGGAGACCTGGATCATGTGAC 914
 Qy 1244 AspArgIleArgValArgGlyAspThrGluCysSerGlyArgValClnIleTrpHis 1263
 Db 913 AACAGATAAGACTTCAGGAAGGCCACTTCTCTCTTGGACGTTGGAGATCTGGCAT 854
 Qy 1264 AlaGlySerTrpGlyThrValCysAspSerTrpAspLeuAlaGluAlaGluVal 1283
 Db 853 GGAGGTTCTCTGGGGACAGTGTGTGATGTTCTGGGACTTGGACGATCTCAGTGGTG 794
 Qy 1284 CysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAlaSerPheGlyGln 1303
 Db 793 TGTCAACACTTGGCTGGTGGTCCAGCTTAAAGCATTTCAAGAGAGAGAGTTGGTCAG 734
 Qy 1304 GlyThrGlyThrIleTrpLeuAspMetArgCysLysGlyAsnGluSerPheLeuTrp 1323
 Db 733 GGGACTGGACCATATGCTCAATGAGTGAAGTGAAGGAATGAGTCTCTCTTGTGG 674
 Qy 1324 AspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyVal 1343
 Db 673 GATTGCTCTGCAGACGCTGGGCCATAGTGTGAGTGTGGCAGCAAGAGACGCTCAGTG 614
 Qy 1344 ArgCysSerGlyGlnSerLeuLys-----SerLeuAsnAlaSerSer 1357
 Db 613 AATTCACAGATATTTCAGTGCAGAAAAACCCACAAAAAGCCACACAGGTCGCTCATCC 554
 Qy 1358 GlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPheIle 1377
 Db 553 CGTCAGTCATCTTTATKGCAGTCGGGATCTTGGGTGTCTCTTGGCCATTTTCGTC 494
 Qy 1378 -----LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgVal 1395

Db 493 GCATTATTCTTCTGACTAAAAAGCGAAGACAGACAGACAGCMS-----CTTGCA GTT 443
 Qy 1396 SerThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeu 1415
 Db 442 CCCCAAGAGAGAGAACTTAGTCCACCAATTCATACCGGAGATGAATCTTGCCTG 383
 Qy 1416 LysArgGluAspProHisGlyThrArgThrSerAspThrProAsnHisGlyCysGlu 1435
 Db 382 AAKGCAGATGATCTGGACCTAATGAATCTCTCAGAAAT-----TCCCATGAGTCAGCT 329
 Qy 1436 AspAlaSerAspThrSerLeuLeuGlyVal-----LeuProAlaSerGluAlaThr 1452
 Db 328 GATTTCAGTCTGCTGGAACATAATTTCTGTGTCTAAATTTCTTCTTATTTCTGGAATGGA 269
 Qy 1453 Lys 1453
 Db 268 AAG 266
 RESULT 11
 BI836459
 LOCUS 603082833F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5221905 5',
 DEFINITION mRNA sequence.
 ACCESSION BI836459
 VERSION BI836459.1 GI:15948009
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 723)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM11558 row: a column: 10
 High quality sequence stop: 708.
 FEATURES
 Location/Qualifiers
 1..723
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5221905"
 /clone_lib="NIH_MGC_120"
 /lab_host="DH10B"
 /note="Organ: pooled pancreas and spleen; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of spleen and pancreas from 28 yo
 male. Library is oligo-dT primed and directionally cloned
 (EcoRV site is destroyed upon cloning). Average insert
 size 1.5 kb, insert size range 1-2.5 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 025. Note: this is a NIH_MGC Library."
 BASE COUNT 173 a 160 c 237 g 153 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.24e-64 Length: 723
 Score: 739.00 Matches: 144
 Percent Similarity: 73.88% Conservative: 37
 Best Local Similarity: 58.78% Mismatches: 54
 Query Match: 9.08% Indels: 10
 DB: 13 Gaps: 0
 US-09-759-130B-381 (1-1453) x BI836459 (1-723)

Db 523 CTCAATGAAGTGAAGTCCACAGGGAATGAGTCTCTCTGTGGGATGCTCTGCGCAGACGC 582

Qy 1330 TrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyVal 1343

Db 583 TGGGGCATAGTGTGGGACACAGGAAGACCTCCAGTG 624

RESULT 13

LOCUS BQ327934

DEFINITION MR4-RT0048-070501-010-g04 RT0048 Homo sapiens cDNA, mRNA sequence.

ACCESSION BQ327934

VERSION BQ327934.1 GI:20945766

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS 1 (bases 1 to 585)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, D.F., de Souza, S.J. and Simpson, A.J.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.J. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR46t2-MR4-RT0048-070501-010-g04&t3=2001-05-07&t4=1>) Seq primer: puc 18 forward High quality sequence start: 24 High quality sequence stop: 584.

FEATURES

source

1. .585

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="RT0048"

/dev_stage="Adult"

/note="Organ: kidney tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESVES PCR (U.S. Letters patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 144 a 122 c 186 g 133 t

ORIGIN

Alignment Scores:

Pred. No.: 2,52e-60 Length: 585

Score: 701.00 Matches: 122

Percent Similarity: 80.83% Conservative: 34

Best Local Similarity: 63.21% Mismatches: 37

Query Match: 8.61% Indels: 1

DB: 14 Gaps: 0

US-09-759-130B-381 (1-1453) x BQ327934 (1-585)

Qy 1128 LysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThr 1147

Db 3 AGGGAGATCGGGAGTTAC-TGCTCAATTCATGCTCTGAGACTGACCACTGAAGCC 61

Qy 1148 GluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerVal 1167

Db 62 AGCAGAGAGCGCTGTCAGGCGCTCTGGAAGTTTTTTTACAAATGAGCTTGGGGCACTGTT 121

Qy 1168 GlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGly 1187

Db 122 GGCAGAGAGTAGCATGCTGAAACCACTGTGGGTGTGGTGTGCAGGCAGCTGGCTGTGCA 181

Qy 1188 GluAsnGluValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpVal 1207

Db 182 GACAAAGGGAATCAACCTTGCATCTTTAGACAAGGCCATGTCATCCATCCCTGTTGGGTG 241

Qy 1208 AspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrp 1227

Db 242 GACAATGTTCAAGTCCAAAGGACCTGACACGCTGTGGCAGTGCCCATCATCTCCATGG 301

Qy 1228 GluArgArgIleSerSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArg 1247

Db 302 CAGAAGAGACTGGCCAGCCCTCGGAGGAGACCTGGATCATGTGCAACAAGATAAGA 361

Qy 1248 ValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrp 1267

Db 362 CTTCAGGAAGACCCACTCTCTGTTCTGGACGTGTGGAGATCTGCATGGAGGTTCTCTGG 421

Qy 1268 GlyThrValCysAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeu 1287

Db 422 GGGACAGTGTGATGACTCTTGGGACTTGGACCATGCTCAGGTGGTGTGCAACAACCT 481

Qy 1288 GlyCysGlySerAlaLeuAlaLeuArgAspAlaSerPheGlyGlnGlyThr 1307

Db 482 GCGTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGTCAGGGACTGGACCG 541

Qy 1308 IleTrpLeuAspMetArgCysLysGlyAsnGluSer 1320

Db 542 ATATGGCTCAATGAAGTGAAGTGCAGAGGAATGAGTCT 580

RESULT 14

LOCUS BE502724

DEFINITION hz78d12.x1 NCI-CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3214103 3' similar to TR:Q07898 Q07898 M130 ANTIGEN PRECURSOR. ;, mRNA sequence.

ACCESSION BE502724

VERSION BE502724.1 GI:9705132

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 653)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov Seq primer: -40UP from Glibco High quality sequence stop: 483.

FEATURES

Location/Qualifiers

1. .653

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3214103"

/clone_lib="NCI-CGAP_Lu24"

/tissue_type="carcinoid"


```
Qy 1245 ArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAla 1264
      ::::::::::: ||| ||| ::::::::::: ||::::::::: ||:::::::::
Db 348 AAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGAGATCTGGCATGGA 289
      ||::::::::: ||::::::::: ||::::::::: ||::::::::: ||:::::::::
Qy 1265 GlySerTrpGlyThrValCysAspSerTrpAspLeuAlaGluAlaGluValValCys 1284
      ||::::::::: ||::::::::: ||::::::::: ||::::::::: ||:::::::::
Db 288 GGTTCCTGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTGGTGTGT 229
      ||::::::::: ||::::::::: ||::::::::: ||::::::::: ||:::::::::
Qy 1285 GlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGlyGlnGly 1304
      ||::::::::: ||::::::::: ||| ||| ::::::::::: ||::::::::: ||
Db 228 CAACAACCTGGCTGTGGTCCAGCTTTGAAGCATTCAAAGAGCAGAGTTTGGTCAGGGG 169
      ||::::::::: ||::::::::: ||::::::::: ||::::::::: ||:::::::::
Qy 1305 ThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAsp 1324
      ||| ||| ||::::::::: ||::::::::: ||::::::::: ||::::::::: ||
Db 168 ACTGACCGGATATGGCTCAATGAAGTGAAGTCAAAAGGAATGAGTCTTCCTTGTGGGAT 109
      ||::::::::: ||::::::::: ||::::::::: ||::::::::: ||:::::::::
Qy 1325 CysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArg 1344
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108 TGTCTGCCAGACGCTGGGGCCATAGTAGTGTGGGCACAAGGAAGACGCTGCAGTGAAT 49
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1345 CysSerGlyGln 1348
      :::::::::::
Db 48 CGCAGGTAG 37
```

Search completed: May 12, 2003, 17:31:28
Job time : 2883 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 12, 2003, 10:35:25 ; Search time 79 Seconds
(without alignments)
5640.523 Million cell updates/sec

Title: US-09-759-130B-381

Perfect score: 8138

Sequence: 1 MMLPQNSWHIDFGRCCHON.....CEDASDTSLGLVLPASEATK 1453

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09759130/runat_06052003_123555_23377/app_query.fasta_1.1607
-DB=Issued_Patents_NA -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blos62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09759130.cgn_1.1.41@runat_06052003_123555_23377 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6 -FCAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2435	29.9	5802	4	US-09-341-587-4
2	1142.5	14.0	4360	1	US-08-470-350B-1
3	785.5	9.7	29598	4	US-09-341-587-6
4	715.5	8.8	2001	4	US-09-341-587-2
5	691.5	8.5	2178	3	US-09-034-916-1
6	634.5	7.8	2262	3	US-09-276-400-3
7	634.5	7.8	2262	4	US-09-448-076-3
8	634.5	7.8	2262	4	US-09-702-572-3
9	634.5	7.8	2920	3	US-09-276-400-1
10	634.5	7.8	2920	4	US-09-448-076-1
11	634.5	7.8	2920	4	US-09-702-572-1
12	618.5	7.6	28720	4	US-09-341-587-7

13	314	3.9	2285	1	US-08-477-674-9	Sequence 9, Appl
14	314	3.9	2285	1	US-08-473-791-9	Sequence 9, Appl
15	314	3.9	2285	2	US-08-316-714-9	Sequence 9, Appl
16	314	3.9	2285	3	US-08-473-673-9	Sequence 9, Appl
17	280.5	3.6	1868	1	US-08-392-367B-1	Sequence 1, Appl
18	290.5	3.6	1868	3	US-08-893-467A-1	Sequence 1, Appl
19	290	3.6	2028	1	US-08-453-117-1	Sequence 1, Appl
20	290	3.6	2028	2	US-08-948-222-1	Sequence 1, Appl
21	290	3.6	2028	2	US-08-973-145-1	Sequence 1, Appl
22	290	3.6	2028	2	PCT-US96-08081-1	Sequence 1, Appl
23	290	3.6	2037	1	US-08-134-365-1	Sequence 1, Appl
24	288.5	3.5	1560	2	US-08-794-795-5	Sequence 5, Appl
25	288.5	3.5	1560	4	US-08-249-200-5	Sequence 5, Appl
26	284.5	3.5	1703	2	US-08-794-795-1	Sequence 1, Appl
27	284.5	3.5	1703	4	US-09-249-200-1	Sequence 1, Appl
28	282	3.5	410	4	US-09-221-298-83	Sequence 83, Appl
29	268	3.3	1588	6	5510466-3	Patent No. 5510466
30	248.5	3.1	9723	1	US-08-083-590A-21	Sequence 21, Appl
31	248.5	3.1	9723	3	US-08-532-384-21	Sequence 21, Appl
32	244	3.0	3955	4	US-09-214-278-4	Sequence 4, Appl
33	243.5	3.0	4315	3	US-08-882-046-3	Sequence 3, Appl
34	243.5	3.0	8257	4	US-09-484-970B-65	Sequence 65, Appl
35	239.5	2.9	4483	3	US-08-611-729A-7	Sequence 7, Appl
36	239	2.9	4464	2	US-08-400-159-7	Sequence 7, Appl
37	208.5	2.6	5804	4	US-09-369-364A-12	Sequence 12, Appl
38	207	2.5	7286	5	PCT-US95-11684-1	Sequence 1, Appl
39	200	2.5	6677	4	US-08-939-366-27	Sequence 27, Appl
40	200	2.5	6677	4	US-09-467-997-6	Sequence 6, Appl
41	196.5	2.4	4394	2	US-08-750-152A-1	Sequence 1, Appl
42	193.5	2.4	16382	4	US-08-718-388-8	Sequence 8, Appl
43	186	2.3	4208	4	US-09-214-278-6	Sequence 6, Appl
44	185	2.3	5499	3	US-08-479-722B-1	Sequence 1, Appl
45	185	2.3	5502	5	PCT-US95-02251-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-09-341-587-4
; Sequence 4, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341.587
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: PCT/DE98/00096
; EARLIER FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 5802
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-341-587-4

Alignment Scores:
Pred. No.: 1.93e-215 Length: 5802
Score: 2435.00 Matches: 533
Percent Similarity: 46.58% Conservative: 155
Best Local Similarity: 39.09% Mismatches: 415
Query Match: 29.92% Indels: 374
DB: 4 Gaps: 28

US-09-759-130B-381 (1-1453) x US-09-341-587-4 (1-5802)

Qy	43	GlyThrAsp-----LeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr	60
Db	389	GGATCTGATCTTGGTTCCTGAGCTGGTGAATGGAGATGGCAGGTGTCTAGGCCCA	448
Qy	61	ValGluValLysPheGlnGlnTrpGlyThrValCysAspGlyTrpAsnThr	80

Db 449 GTGGAGATCCTATACAGAGCTCTCGGGCCACCGTGTGTGTGATGACAGCTGGGACACCAAT 508
Qy 81 AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMet----- 97
Db 509 GATGCCAACGTGGTGTAGGACAGTGGTGTGGCTGGCCATGTCACGCTCCAGGAAT 568
Qy 98 PheArgPheGlyGlnAlaValThrArgHisGlyLysIleThrLeuAspValSerCys 117
Db 569 GCCTGGTTTGGCCAGGC-----TCAGGACCCATTGGCTGTGATGATGGGCTGC 619
Qy 118 TyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpCysSerHisAsnCys 137
Db 620 TCAGGACACCAATCTTACCTGTGGAGCTGCCCCACAATGGCTGGCTCTCCCACTACTGT 679
Qy 138 TyrHisGlyGluAspValGlyValAsnCysTyr----- 148
Db 680 GGCATGGTGAAGTCTGGTGTATCTGCTCAGCTGCCCGCCAGCTCAGTCAACACTCAGG 739
Qy 149 -----GlyGluAlaAsn 152
Db 740 CCAGAAAGTTGGCTGTCCAGGATATCACCACCTGTACCCACAGAGGATCTGAATCCAGT 799
Qy 153 LeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArgValGluValLysPhe 172
Db 800 TTGGCCCTCAGGCTGGTGAATGAGCGCAGAGTGTGAGGCGGAGTGGAGTTCCTATAC 859
Qy 173 GlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValVal 192
Db 860 CGAGGCTCTGGGACCGTGTGTGATGACTACTGGGACACCAATGATGCCATGTGGTTC 919
Qy 193 CysArgGlnLeuGlyCysProSerPheIleSerSerGlyValValAsnSerProAla 212
Db 920 TGCAGGACGCTGGCTGTGGCTGGCCATGTCAGGCCAGGAATGCCAGTTTGGCCAG 979
Qy 213 ValLeuArgProIleTrpLeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrp 232
Db 980 GGCTCAGGACCCCATTTCTGGATGATGTCCGCTGTCTCAGGACGAGTCTACCTGTGG 1039
Qy 233 AsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHisAsnGluAspValThrLeu 252
Db 1040 AGTGTCCCAACAATGGCTGGCTCACCCACAACCTGTGGCCATAGTGAAGACGCTGGTGC 1099
Qy 253 ThrCysTyr----- 255
Db 1100 ATCTGCTAGCTCCCGAGTCCCGCCGACACCCAGCCAGATCTTGGCCGACCTCACAT 1159
Qy 256 -----AspSerSerAspLeuGluLeuArgLeuValGlyThrAsnArg 270
Db 1160 GCATCAACAGCAGGACCTGAATCCAGTTTGGCCCTCAGGCTGGTGAATGGAGTGACAGG 1219
Qy 271 CysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyThrValCysHisLys 290
Db 1220 TGTCCAGGCGGAGTGGAGTCTATACCGAGGCTCTCGGGCACCGCTGTGTGATGATAGC 1279
Qy 291 TrpAsnAsnAlaAlaAspValValCysLysGlnLeuGlyCysGlyThrAlaLeuHis 310
Db 1280 TGGGACACCAAGTACGCCCAATGTGGTCTCCGCGCAGCTGGCTGTGGCTGGGCGACGTCA 1339
Qy 311 PheAlaGlyLeuProHisLeuGlnSerGlySerAspValValTrpLeuAspGlyValSer 330
Db 1340 GCCCCAGGAATGCCCGTTTGGCCAGGTTTCAGGACCAATGCTCTGGATGACGTGGCG 1399
Qy 331 CysSerGlyAsnGluSerPheLeuTrpAspCysArgHisSerGlyThrValAsnAsp 350
Db 1400 TGCTCAGGCTATGAGTCTACCTGTGGAGCTGCCCCACAATGGCTGGCTCTCCCAATAAC 1459
Qy 351 CysLeuHisGlnAsnAspValSerValIleCysSerAspGlyAlaAspLeuGluLeuArg 370
Db 1460 TGTACGACAGTGAACAGCGTGTGTCTATCTGCTCAGCTGCC----- 1501
Qy 371 LeuAlaAspGlySerAsnAsnCysSerGlyArgValGluValArgIleHisGluGlnTrp 390
Db 1502 -----CACTCCTGG 1510

Qy 391 TrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeu 410
Db 1511 TCGAGC----- 1516
Qy 411 GlyCysProPheSerValPheGlySerArgAlaLysProSerAsnGluAlaArgAsp 430
Db 1517 -----CCAGTCCAGACACATGGCG 1537
Qy 431 IleTrpIleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyr 450
Db 1538 ACATCACCCTTCCCTGCATCGACAGTAGGATCTGAATCCAGT----- 1579
Qy 451 AspGlyLysAlaLysArgThrCysPheArgArgSerAspAlaGlyValIleCysSerAsp 470
Db 1579 ----- 1579
Qy 471 LysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTyrGlyArgLeuGlu 490
Db 1580 -----TTGCCCTGAGGCTGTAATGGAGTGCACAGGTGCAGGGCCGAGTGGAG 1630
Qy 491 ValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSerThrArgAsnAla 510
Db 1631 GTCTATACCAAGCTCTCGGGCACCGTGTGGATGACAGCTGGGACACCAATGATGCC 1690
Qy 511 AlaValValCysLysGlnLeuGlyCysGlyLysProMetHisValPheGlyMetThrTyr 530
Db 1691 AATGCTGTGCAGGCAACCGGCTGTGGCTGGCCATGTGCAGCCCGAGGAATGCCCG 1750
Qy 531 PheLysGluAlaSerGlyProIleTrpLeuAspAspValSerCysIleGlyAsnGluSer 550
Db 1751 TTTGGTCAAGGCTCAGGACCCCATTTCTGGATGATGTGGCTGCAGGACACGAGTCT 1810
Qy 551 AsnIleTrpAspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgGluAsp 570
Db 1811 TACCCGTGAGCTGCCCCACCAATGGCTGCTCTCCCAACACTGTGCCATAGTGAAGAC 1870
Qy 571 ValIleValThrCysSerGlyAspAla-----ThrTrp--- 581
Db 1871 GCTGGTGTCTCTGCTCAGCTTCCAGTCCCGGCCAACACCTAGTCCAGACACTTGGCCA 1930
Qy 582 -----GlyLeuArgLeuValGly 588
Db 1931 ACCTCATGCTATCAACAGCAGGATCTGAATCAGTTTGGCCCTGAGGCTGGTGAATGA 1990
Qy 589 SerAsnArgCysSerGlyArgLeuGluValThrPheGlnGlyArgTrpGlyThrValCys 608
Db 1991 GGTGACAGGTGTCCAGGCGGAGTGGAGTCTTATACCGAGGCTCTCGGGCACCGTGTGT 2050
Qy 609 AspAspGlyTrpAsnSerLysAlaAlaValValCysSerGlnLeuAspCysProSer 628
Db 2051 GATGACTACTGGGACACCAATGATGCTGTTTGCAGGACGCTGGCTGTGGCTGG 2110
Qy 629 SerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyLysIleTrpLeuAsp 648
Db 2111 GCCATGACCCCGGAGAAATGCCGTTTGGCCAGGTTTCAGGACCAATGCTCTGGAT 2170
Qy 649 AspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgAsnSerGlyTrpGly 668
Db 2171 GATGTGCGTGTCTCAGGACATGATGCTATCTGTGGAGTGTGCCCGCCACAATGGCTG 2230
Qy 669 AsnAsnAspCysSerHisSerGluAspValGlyValIleCysSer----- 683
Db 2231 TCCCAACAATGTGGCCATCATGAAGACGCTGGTGTCTATCTGCTCAGCTTCCCAAGTCC 2290
Qy 684 -----AspAlaSer 686
Db 2291 CCGACACCCAGCCGACACTTGGCCCACTCATCATCATCAACAGCAGGATCTGAATCC 2350
Qy 687 AspMetGluLeuArgLeuValGlySerSerArgCysAlaGlyLysValGluValAsn 706
Db 2351 AGTTTGGCCCTGAGGCTGGTGAATGGAGTGTGACAGGTTGTGAGGCGGCGGAGTGGAGT 2410

Qy 707 ValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsnIleAlaGluVal 726
Db 2411 TACCGAGGCTCTGGGCGACCGTGTGTGATGACTACTGGGACACCAATGATCCCAATGTG 2470
Qy 727 ValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGluProHisPheThr 746
Db 2471 GTTTCAGGACGCTGGGCTGGGCTGGGCGCCAGCTAGCCCGGAAATGCGCGGTTGGC 2530
Qy 747 GluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGlyGluAlaSerLeu 766
Db 2531 CAGGTTTCAGGACCACTGTCTCGATGATGCTGCTCAGGACATGATCTATCTG 2590
Qy 767 TrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsnMetGluAlaSer 786
Db 2591 TGGAGCTGCCCGCCCAATGGCTGCTCCCAACAACCTGTGCCATCATGAAGACGCTGCT 2650
Qy 787 LeuIleCysSerAlaHisArg---GlnProArgLeuValGlyAlaAspMetProCysSer 805
Db 2651 GTCATCTGCTCAGCTTCCAGTCCCGCGG--- 2680
Qy 806 GlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSerAspPheSer 825
Db 2681 -----ACACCCAGCCAGACACTTGGCCAACT----- 2707
Qy 826 LeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIleSerLeuSer 845
Db 2707 ----- 2707
Qy 846 ValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPheGlnCysGlu 865
Db 2708 -----TCTCGTGCATCAACAGCA 2725
Qy 866 GlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAspThrCysIle 885
Db 2726 GGATCTGAATCACTTTGGCCCTG----- 2749
Qy 886 HisSerArgGluValGlyValValCysSerArgTyrThrAspValArgLeuValAsnGly 905
Db 2750 -----AGACTGCTGAATGCA 2764
Qy 906 LysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCys 925
Db 2765 GGTGACAGGTGTGAGGCGGAGTGAGTCTCTATACCAAGGCTCTCTGGGCGACCGTGTGT 2824
Qy 926 AspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSerCysGlyThr 945
Db 2825 GATGACTACTGGGACCAATGATGCCAACCTGTGTCTCAGCAGCAGCTGGGCTGTGGCTGG 2884
Qy 946 AlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArgValTrpGlyHis 965
Db 2885 GCCATGTGAGCCCGCCAGGAAATGCCAGTTTGGCCAGGCTCAGGACCCATTTCTCTGGAT 2944
Qy 966 ArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThrValLeuGly 985
Db 2945 GATGTGCGCTGCTCAGGACAGAGTCTTACCTGTGGAGCTGCCCCCAATGGCTGGCTC 3004
Qy 986 AlaProProCysIleHisGlyAsnThrValSerValIleCysThrGlySerLeuThrGln 1005
Db 3005 TCCCAACAATGTGGCCATCATGAAGATGCTGTGTCTATCTGCTCAGCTGCTCAGTCCAG 3064
Qy 1006 -----ProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaValPro 1023
Db 3065 TCAACGCCCGCCAGAGATCTGGCTGACCAACCACTTACCGGCATTCACAGTAGTAGTCT 3124
Qy 1024 GluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAspSer 1043
Db 3125 GAATCCAGTTTG-----GCTCTGAGGCTGGTGAATGGAGGTGAC 3163
Qy 1044 ArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAspAsp 1063
Db 3164 AGGTGTCCAGGCCGAGTGGAGGTCTGTATCGAGCTCTCTGGGACCGCTGTGTGATGAC 3223
Qy 1064 GlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAlaPhe 1083

Db 3224 AGCTGGGACACCAATGATGCCAAATGTGCTCTGCAGGAGCTGGGTGGTGGCCATG 3283
Qy 1084 AsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspLeu 1103
Db 3284 TCGGCCCCCAGGAATGCGCGTGTGGCCAGGCTCAGGACCATTTGCTGTGATGATG 3343
Qy 1104 AsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHis 1123
Db 3344 CGCTGCTCAGGAAATGACTCTACCTGTGGAGCTGCCCCACAAAGGCTGGCTCACCCAC 3403
Qy 1124 AspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeu 1143
Db 3404 AACTGTGCCCATCAGAAAGCGCTGTGTCTCTCTCA-----GCCACCCAAATA 3454
Qy 1144 TyrSerGluThr-----GluThrGlu 1150
Db 3455 AATCTACTACAGAGATGGTGGCATCCAACTACACCACTGCAAGACCCCTCTTCA 3514
Qy 1151 SerCysAlaGlyArgLeuGluValPheTyr---AsnGlyThrTrpGlySer 1166
Db 3515 AATGTGTGGCTTCTTA-----TTCTATGCCAGTGGACATTTCTCCAGCCCATCTAC 3568
Qy 1166 ----- 1166
Db 3569 CTGTGATACTACCCCAACAATGCTAAGTGTGTTGGGAAATAGAAAGTGAATTTCTGTTAT 3628
Qy 1167 -----ValGlyArgArgAsnIle----- 1172
Db 3629 CGCATAAACCTGGGCTTCTCAGTAACTGAAATTTGGAGGACACCACTAATCGCATTTGAT 3688
Qy 1173 -----ThrThrAlaIleAlaGlyIleValCys--- 1181
Db 3689 TATGTTGAAATCTTTGATGGATCATTAATAGCAGCTCTCTCTGGGAAATCTGTAAT 3748
Qy 1182 -----ArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeu-----AlaPro 1196
Db 3749 GATACCGAGCAATATTATACATCTTCTTACAACCAAGAAATGACCATTCACCTTCGAAGTAC 3808
Qy 1197 LeuSerLysThrGlySerGlyPheMet---TrpValAspAspIleGlnCysProLysThr 1215
Db 3809 ATCAGTTTCCAAAACACACCTGGCTTTTGGCTGGTGTATAC----- 3847
Qy 1216 HisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerProAla 1235
Db 3848 -----TCTTCCCAAGC 3859
Qy 1236 GluGluThrTrpIleThrCysGluAspArgIleArgValArgGlyAspThrGluCys 1255
Db 3860 GATGCCACCTTG-----AGGTGGTCAATTTTAAATTCATCTATGCTATGT 3907
Qy 1256 SerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThrValCysAspSerTrp 1275
Db 3908 CCGGCGGTGTAGAATAATTTACCATGTGGCAGCTGGGGACAGTTGTGATGACTCTGG 3967
Qy 1276 AspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAla 1295
Db 3968 ACCATTACAGGAAGCTGAGTGTCTGCAGACAGCTAGGTGTGGAGCTGCAGTTTCAGCC 4027
Qy 1296 LeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrpLeuAspMetArgCys 1315
Db 4028 CTTGAAATGATATATTTGGCTCTGGCTGCGCCCATCCACCTGGACGATAGATGTC 4087
Qy 1316 LysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCys 1335
Db 4088 TCAGGAGCGGAATCCACTCTCTGGCAGTGGCGGAACCGAGGCTGTTCTCCACACATGT 4147
Qy 1336 GlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSerLeuLysSer 1352
Db 4148 AATCATCTGAGATGCTGGTGTCTATCTGCTCAGGAAACCATCTATCGACA 4198

RESULT 2

US-08-470-350B-1

; EARLIER FILING DATE: 1998-01-09
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 29598
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-341-587-6

Alignment Scores:

Pred. No.: 3,89e-61 Length: 29598
 Score: 785.50 Matches: 501
 Percent Similarity: 26.4% Conservative: 185
 Best Local Similarity: 19.3% Mismatches: 520
 Query Match: 9.65% Indels: 1390
 DB: 4 Gaps: 103

US-09-759-130b-381 (1-1453) x US-09-341-587-6 (1-29598)

QY 35 CysPheLeuIleSerSerPheAsnGlyThrAspLeuGluLeuValAsnGlyAsp 54
 DB 3837 TGCCCTCCCTAGGATCTGATCTGGT-----TTGGCCCTCAGGCTGGTGAATGAGAT 3890
 QY 55 GlyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAsp 74
 DB 3891 GGCAGGTGCAGGGCGGAGTGGAGATCCTATACCGAGGCTCCTGGGACCGCTGTGTGAT 3950
 QY 75 AspGlyTrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSer 94
 DB 3951 GACAGCTGGGACCAACATGATCCCAACGTGGTCTGTAGGCAGCTGGTGTGGCTGGGCC 4010
 QY 95 PheAlaMet-----PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrp 111
 DB 4011 ATGTGAGCTCCAGGAATACCTGGTGTGGCCAGGCG-----TCAGGACCCATGTGCC 4061
 QY 112 LeuAspAspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGlu 131
 DB 4062 CTGGATGATGTCGGTGTCTCAGGACACGATCTTACCTGTGGAGCTGCCCCCAACATGCC 4121
 QY 132 TrpGlySerHisAsnCysTyrHisGlyGluAspValGlyValAsnCysTyrGlyGluAla 151
 DB 4122 TGCTCTCCCAATACGTGGCCATGCTGAAGATGCTGTGTGTATCTGCTCAGGTAGG--- 4178
 QY 152 AsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCys----- 164
 DB 4179 CATCCAGATCTCTGGAGGTTGGTGTGGTGGCTCATGCTTTAATCCCCACACTTTGGG 4238
 QY 165 -----SerGlyArgValGluValLysPheGlnGluArgTrp 176
 DB 4239 AGGATGAGTAGGCAGATGCTGTAGCTCAGGCGTTCAAGACCACTCTAGGCCAAGATGG 4298
 QY 177 GlyThr----- 178
 DB 4299 CAACCCCATCTCTATTAAAAAGAAAATCACTTTGGGCCCATTTCTGGGGACAGACTGTA 4358
 QY 178 ----- 178
 DB 4359 TTCAGAGATATAGTAACAGGTTTCTCACTCCAGGGCCCCGCCCAAGTCTTTCTCGTGT 4418
 QY 179 -----IleCysAspAspGlyTrp 184
 DB 4419 ATCTGTGGCTCTGAGTTTAGTTGGGCGAGTTGGCCCTGTGGGTACATGCCACGGTCAG 4478
 QY 185 AsnLeuAsnThrAlaAlaValValCysArgGlnLeuGly----- 197
 DB 4479 CACAATCCCAAGTCGCGAGCAGTGCCGCCAACATTTGTTGGAGGTGACTGAGGCAACA 4538
 QY 197 ----- 197
 DB 4539*GTGTTACACACAGGGTCCAGCATGCTGACATCAGAAATGATGACCTTTTATCTGCTC 4598
 QY 197 ----- 197

DB 4599 CAGGAACATCCGCATAACTCATACCCCATTTTCCCGAGGCTCTCTCAGGGACAGCA 4658
 QY 198 -----CysProSerSerPhe 202
 DB 4659 CGTTTTTGGGATCTGGTGGTGAAGCCCTTCCCTCCCTTGAACACTGTATTACCTTGAGT 4718
 QY 203 IleSerSer-----GlyValValAsnSerProAlaValLeu--- 214
 DB 4719 GTTGTGACGCCGCTGACCCAGAGGAAGGATGCAATTGTAAATAGTTCACTTATGATTGCT 4778
 QY 215 -----ArgProIle-----TrpLeuAspAspIleLeuCysGlnGlyAsn 227
 DB 4779 GTGAAGAGAGGAGGAGAGTGGGTAAAGGTGGGAGGAT-----GSCAT 4823
 QY 228 GluLeuAlaLeu-TipAsn-----CysArgHisArgGlyTrpGlyAsnHisAspCys 244
 DB 4824 GAGTCGCAATGCTGGGCTCACAGTACAGTACAGTACAGGAGGAGGAGGAGCTCTCTAGA 4883
 QY 244 sSerHisAsnGluAsp----- 249
 DB 4884 GTATCACCAGGAATCCAGCTCAACCTCAGACTCATCGCAGCTCAAGCGTGGAGAGTTAG 4943
 QY 250 -----ValThrLeuThrCysTyrAspSerSerAspLeuGlu 261
 DB 4944 GGTCTTCATCTACTCATGCTTAGGGTGGCTGTTTCTGCTCTCCCTGGGTGTTGATACCGA 5003
 QY 261 uLeuArgLeuValGly----- 266
 DB 5004 GGTGTTGTGGCGGTCCAGGACCCGAGTCACTCATATCCCTGGGAGAGTGAGCCAC 5063
 QY 267 -GlyThrAsnArgCysMet----- 272
 DB 5064 AGGCACAGAGAAGCAATGCCAGGGGCTGTGGGGTCCCTGAGGGACCCGAGGCTTCACGG 5123
 QY 273 -----GlyArgValGlu-LeuLysIleGlnGly----- 281
 DB 5124 TGGGCACCTGGCAGGGCCCACTTGGTGGCGTGTGATGGGCATCAGCTCAGGGGTGATAGATA 5183
 QY 281 ----- 281
 DB 5184 CCCCAGTCACTTTCAGCCTTAACCTACTTGGAGTCACTGAGTGTGTGGTGTCTAATGTT 5243
 QY 281 ----- 281
 DB 5244 GCTATTTTTTCTCAGAGTGGCCAGCTCAGTCAACACTCAGGCCAGGTGAGTCCCCCAG 5303
 QY 281 ----- 281
 DB 5304 AATCCTTCTCGGGATACCCCTTCTTCTGCTCAGTTACCCCTTCCCTACTCCACAGA 5363
 QY 282 -----ArgTrpGlyThrValCysHis----- 288
 DB 5364 GGCCTCTCTCTCTGCGAGACTCTGGGCAATATTTTCCACCCCACTCTGTAAC 5423
 QY 289 -----HisLysTrpAsnAlaAlaAla----- 296
 DB 5424 GAGACCTTAGCATGGGGTCTTTAACACACATGGGATTTAGCTTCTGCTTCTCTCAG 5483
 QY 297 -----AspValValCys----- 300
 DB 5484 TCCATCTCAGCTTTCATCAGCAGTTCCTGTCCTCCCAATGAACAACTCTTAGAGTTC 5543
 QY 301 -LysGlnLeuGlyCys-----GlyThrAlaLeuHis-PheAlaGlyLeu----- 314
 DB 5544 AGAAGAAGACCTGCTGTTTAATGAGCTTTGGCCTCTTTATATTAGGGCTGACACAACCT 5603
 QY 314 ----- 314
 DB 5604 TTCGGGGAATTTGAGGTTGAGACTCTCTAATGTTCTGTTGAAGGCAAGGTCACTACAGCT 5663
 QY 315 -----ProHisLeu----- 317
 DB 5664 TGATACCCCATCTCTCACTACTTGAACCAACCCAGATTTCTGCAGGGCCACGCTCGCAT 5723

D	b	7840	TTGCGTGGGTGGAGTTCTTTGACCTCAGCTCTTCTCAGAACCCTGCTGAGCATTCGCCCTG				
Q	y	682	CysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGly		:::
D	b	7900	TGTTTAGGTCTGGTGAG--		:::	GGG
Q	y	702	LysValcIuValasnValgInGlyAlaValGlyIleLeuCysAlaAasnGlyTrpGlyMet	:::					721
D	b	7921	AGGCAGCTCCCATGAGGCCACGACATGSCCTTGTCATTGCTGTGATTGGGCGCTG						7980
Q	y	722	AsnIleAlaGlu-----ValValCysArgGlnLeuGluCysGlySerAla						736
D	b	7981	AAGATCGCAACAAGGGATTTTGGCTGGAGTGGCTCTCTCAGCCITTCGTACTCAGGAACAC						8040
Q	y	737	IleArgValSerArgIu-----ProHisPheThrGluArgThrLeu						750
D	b	8041	CTAAGATGTGCRAAGGAGTGGTGGTTAGGTCAACGGGTTCCTCGCCGGCAGACACAA						8100
Q	y	751	-----HisIleLeuMetSerAsnSergLy-----CysThr-GlyGl						762
D	b	8101	TTTGTATCACTCAGAGCTGGCAATAGTGACAGGATCTGCCTCACGCCCTTACACGCT--						8158
Q	y	762	yGluAlaSerLeuTrpAspCysIleArgTrpCluTrpLysGlnThrAlaCysHISLeuas						782
D	b	8159	--GCATCTCTGTGGGGATG---GCATGGCAATGT---CCCTCCCTGTGTAGTAGGAA						8208
Q	y	782	nMetGlu-----Alase						786
D	b	8209	CTAGGATGGACTGAGTGTCAAGACTCGCTCATTTCTTCCCTCCTCGTTCAGATTTTCCG						8268
Q	y	786	rLeuIleCysSer-----						790
D	b	8269	ACTTCTGTGTAAATGTTCTCGATCTGACCTTCTTCTTCTTCTCACAGCTCCCAGTCC						8328
Q	y	791	-AlaHisArgGlnProArgLeuValcIuAlaaspMetPro-----						803
D	b	8329	GGCGGACCCAGCCCCAGGTAGTCCCGAGTGTCTTCTCTCAAATGTCCCTTCTCTTC						8388
Q	y	803	-----						803
D	b	8389	TGCCAATCACCCTTCCACACTCCACAGAGCTTCCTGTTTTCTGTGTGATACTGTG						8448
Q	y	803	-----						803
D	b	8449	GGGCATATTATTTCCACCCCCAACACCGCGTGTGTACTGAGATCCCGACAGCGCTTT						8508
Q	y	803	-----						803
D	b	8509	TTAACACACACAGGATTGGAGGGCCTCTGCTTCTTTTCAACCCCTCTCAGCTTTTCAT						8568
Q	y	803	-----						803
D	b	8569	GAAACAGTTTTCATACTGTCCAGTGGCAACCCCTTACAGGTTCAGGAAGTGGCCCCATGT						8628
Q	y	803	-----						803
D	b	8629	TTAATGAGCTTTGGTCCCTTTTATTTCCGAGCTCACATATAGTTTCTGAAAATTGAGGGT						8688
Q	y	804	-----CysSerGlyArgValGluValLysHis-----						812
D	b	8689	GTCACTCTGACCCCGTTCAAGGCATCTCAGGGCAGGCTCGATACCCCATCCATCACT						8748
Q	y	813	-AlaAspThrTrpArgSerValCysAspSerAspPheSerLeuHisAlaLaAsnValle						832
D	b	8749	GCTGGAACAATCCGAGATTAT---GATTCCGAGATTCTATCCAGAGAGGCAGAGTTT						8805
Q	y	832	uCys-----ArgGluLeu---AsnCysGlyAspAlaIleSerLeu-----						844
D	b	8806	GTGCTCGGCAGGAGAGAGGGGATAATAAAGGTTT--GTGATGCTCTGCTTAACAGAAACC						8864
Q	y	844	-----						844

Db	8865	TGATTCTCTGATTGCTCCCTGGGAGCCCTGGTTCCCTTAACATTTTACCTGGCAGTGTCT	8922
Qy	845	-----SerValGlyAs	848
Db	8925	CGAGCTTCACCAATGGCGTCTGATGCTCTAGTCACTGCATCATCAGCAGATGTGGGATG	8984
Qy	848	pHisPheGlyLysGlyAsnGlyLeu	856
Db	8985	GCAT-----GGTGGGGCATCTCTAACAGATAGAAAGATACCCAGATTACAGA	9035
Qy	856	-----	856
Db	9036	AATGGAGGGCTCAGTCTGTCTCCAGCAGACCTTTGTCCCTGGATGAGTTTCACAGCACA	9095
Qy	857	----ThrTrpAlaGlyLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCys-p	875
Db	9096	GGAGACCTGGGAACAGACACATGGGAACAATGGCAGGAACGAAGTGGAAATCTTGTC	9155
Qy	875	rolleValGlnHisProGluAsp	882
Db	9156	CGTTGATTCTCTCTCCAGAGATCCTTTTGTCTCTGGCTTTTCCCTTCAAGTCTAATTC	9215
Qy	883	-----ThrcysIleHisSerA	888
Db	9216	TGTCCTTTTCTTTTGTGTGCAATTTACAGATACTTGGCGACCTCACAATGCATCAACAGCA	9275
Qy	888	rgGluVal-----GlyValValCysSerArgTyr	897
Db	9276	GGTAATAACCCCTCTCACCCCTCCCTAGGACTCACTATCTCTGGACATATTTTGTGTTTG	9335
Qy	898	--ThrAsp-----ValArgLeuValAsnGlyLysSer	907
Db	9336	AACTGATAGGATGAGCTCAATGTGGCTTCTCTGTTTTCATGTCCCTGTGGTTCGCT	9395
Qy	908	--GlnCysAspGlyGlnValGluIleAsnValLeuGly	919
Db	9396	GGGAGGAAGTGGAAATCTCTGAGGAGCCAGTCTCTGGGTCTGATGTTTGAGGACGGAGGT	9455
Qy	920	-----HisTrpG	922
Db	9456	GTTCGGTCACCTGTCTCCATGGAACTCTGTTCCAAAGTGGTCAGGAAGAATCCTCATCCAG	9515
Qy	922	lySerLeuCysAspThrHisTrp	931
Db	9516	GTGCTCAG--GACGAGCACTGGAGGGCTCTTAATGCTGTGGAACTCATCTCTGGCC	9572
Qy	931	roGlu-----AspAlaArgValLeu	943
Db	9573	CTCAGGCATGGGATCCAGACCTCTCAAGACGGGTGAAAGTGCCTGTCCCTGCACCTGT	9632
Qy	943	ysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGly-GluArgSerValArgVal	962
Db	9633	GTGCCAAGCCCTTGCCATCATTGGCAATTTGCCAGAGGCAGAGAGGCCCATGCAGGTG	9692
Qy	963	-----trpGlyHisArgPheHisCysLeuGlyAsnGlySerLeuGly	976
Db	9693	CCAATAAGCTCCTGAATATGG-----AGGGGTCTCAGGCCTGTCTATT	9734
Qy	977	AspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHis	991
Db	9735	TGTAGCTGTGTA-----GTCCATCTCTGTGTTTCCACCCAGAGTGGGGAGTG	9779
Qy	992	-----GlyAsnThrValSerValIleCysThrGlySerLeuThr	1004
Db	9780	GGGTGTCATTCTCTCTCTGGGGTCATACATGCTTACCACATAGCTTAGCTTAGCTTT	9839
Qy	1005	-----GlnProLeuProCysLeuAlaAsnValSerAsp	1016
Db	9840	TATAGACTTGAGTAGAATAGGGCATCACITTTTCCACATATGACCAAGCTTAACCTCTGG	9899
Qy	1016	-----	1016
Db	9900	TGCAGGCATCTGCATTTGACTGCGTGGCCCTGGTACCTTTGGCATTTGTATTCAAACCT	9959

```
Qy 1017 ProTyrLeuSerAlaValProGlu----- 1024
Db 9960 GACFACCTTTCCCATGTCCTGAGTGTGTCCTCATGCCCTTTCCCTTCACCTCAGATGGAG 10019
Qy 1025 ---GlySerAlaLeuIleCysLeu-----GluAspLysArgLeuArgLeu 1038
Db 10020 ATGATGAAGGATTTCTTGTGTTCCCTCTAGGACCTGAATCCAGTTTGGCCCTCAGGCTG 10079
Qy 1039 ValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGly 1058
Db 10080 GTGAATGAGGTGACAGGTGACGGCGAGTGGAGGTCCATACCGAGGTCCCTGGGGC 10139
Qy 1059 ThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGly 1078
Db 10140 ACCGTGTGTGATGATGCTGGACACACAGTCAGCCATGTGCTGTCCGGCAGCTGGGC 10199
Qy 1079 CysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIle 1098
Db 10200 TGTGGCTGGCCACGTCAGCCCAAGAAATGCCGGTTTGGCCAGGTTTCAGGACCCATT 10259
Qy 1099 TrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArg 1118
Db 10260 GTCCGTGATGACGTGCGCTGCTCAGGCTATGAGTCTACCTGTGGAGCTGCCCCACCAAT 10319
Qy 1119 GlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSer----- 1136
Db 10320 GGCTGGCTCTCCATAACTGTGACGACAGTGAAGACGCTGGTGTCTATCTGCTCAGGTGG 10379
Qy 1136 ----- 1136
Db 10380 CTTCCAGACCTTGGGCTCCCTCTCTTGGGTGAGATTTTGTCCAGAAAGCGAGGTCTCAT 10439
Qy 1137 -----GluPheThrAlaLeuArgLeuTyr----- 1144
Db 10440 TATGTTCTGATCTCTCACTCAGAGCTTTTTCAGGCTTTCTATATATATCTGATATCTCT 10499
Qy 1145 -----SerGluThrGluThrGluSer-Cys----- 1152
Db 10500 TAGCTCTCTCTAGGAAACTGCATGAGTCTTCATTTGCCAGGTTTTCAGGAGTCCAGGTAG 10559
Qy 1153 -----AlaGlyArgLeuGluVal-----PheTyrAsnGlyThrTrpGln 1165
Db 10560 GACAACGGCGGCAAGATGAATGAAGGTTCACGCCCTTTGTTCACCTACCGAGGACGAGC---GC 10616
Qy 1165 ySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGln 1185
Db 10617 AAGCAGAGGAGAGAGAGAAAGTCCAGGCTCTTTGCC-----TTTAA 10658
Qy 1185 yCysGlyGluAsnGlyValSer-----LeuAlaProLeuSerLysThrGln 1201
Db 10659 CTGTGGCTGGAAGCAATAGTGGGCTGGTGTGTCATGAGAAAGAGGCTGATTTGGG 10718
Qy 1201 ySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGln 1221
Db 10719 TCTTGGCTTTGCA-----GG 10733
Qy 1221 nCysLeuSerAlaProTrpGluArgArgIleSerSer----- 1233
Db 10734 CTGCATA---GGACCATGGTCGACGCGTCTGCTCAGCTTCTGATGAGGCGCTCGGGCTGC 10790
Qy 1234 -----ProAlaGluGluThrTrp-----IleThrCysGluAspArgIleArgVa 1248
Db 10791 TTGTACTCTAGCAAGAGGATGGGAGGTGGCTGGCTGGCAGAGGTCAACAAGGTGAGGGA 10850
Qy 1248 lArgGlyGlyAspThrGluCysSerGly----- 1257
Db 10851 GGAAGGAAGAGAGACCAAGAGGAGGCTCAGACTCTTCAACCACCACTCTTGTCTAGGA 10910
Qy 1257 ----- 1257
Db 10911 ACTAAGAAGAAGACTCACCCCTGACCGAGGGGACCTAAGTTATTCATGACGAGTGTGGT 10970
```

```
Qy 1258 ---ArgValGluIleTrp-----HisAlaGlySerTrpGlyThrValCysAs 1272
Db 10971 CTCCATGACCCAGACCTGGTGCATCAGGCTCACCTCTAAACTTGGGATTCAGTTGCCAA 11030
Qy 1272 pAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCys----- 1289
Db 11031 TATAACA-----CTTGGATGTGACAAGCC 11054
Qy 1290 -----GlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300
Db 11055 TCTAAAGTATAGCACCCCTGGCTCTCCAGGTTCCATCTTCCCTACTAGCAAGAT----- 11109
Qy 1300 rPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSe 1320
Db 11110 -TTTCAAAAGGT-----TACATGGATGAA----- 11133
Qy 1320 rPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluSe 1340
Db 11134 -----GCACATATATAGGA 11147
Qy 1340 pAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLe 1360
Db 11148 TGGGGA-----GGGACTCTCTCTCTGGGAGGAGCCCAAGACACAG 11189
Qy 1360 uAlaLeuIleLeuSerSerIlePhe 1368
Db 11190 GCTCTCTAGTAGAAGCGGCTCTC 11214
RESULT 4
US-09-341-587-2
; Sequence 2, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341.587
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: PCT/DE98/00096
; EARLIER FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-341-587-2
Alignment Scores:
Pred. No.: 1,51e-56 Length: 2001
Score: 715.50 Matches: 164
Percent Similarity: 48.18% Conservative: 61
Best Local Similarity: 35.12% Mismatches: 119
Query Match: 8.79% Indels: 123
DB: 4 Gaps: 15
US-09-759-130B-381 (1-1453) x US-09-341-587-2 (1-2001)
Qy 956 GlyGluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeu 975
Db 1 GGACACGAGTCTTANABCTGTGGAGC-----TGCCCCCAACAATGGCTGGCTC 48
Qy 976 LeuAspAsnCysGlnMetThrValLeuGlyAlaProCysIleHisGlyAsnThrVal 995
Db 49 TCCCAACAATGGCTGGC-----CATCATGAAGATGCT 78
Qy 996 SerValIleCysThrGlySerLeuThrGln-----ProLeuPheProCysLeuAlaAsn 1013
Db 79 GGTGTCTATCTGCTAGCTGCTCAGTCCAGTCAAGCCAGCCAGGATACCTTGGCTGACC 138
Qy 1014 ValSerAspProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAsp 1033
Db 139 ACCAACTTACCGGCATTGACAGTAGGATCTGAATCCAGTTTG----- 180
```

QY 1034 LysArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluLeuTyrHis 1053
|||||
Db 181 ---GCTCTGAGGTGCTGAATGAGGTGAGTGTCTGAGGCGGAGTGTGAGTCTGTAT 237
QY 1054 AspGlyPheThrPheThrLeuValAspGlyTyrPheSerLeuSerAspAlaHisValVal 1073
|||||
Db 238 CGAGGCTCTGGGACCGCTGTGTGATGACAGCTGGGACCAATGATGCCAATGTGGTC 297
QY 1074 CysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGlu 1093
|||||
Db 298 TCGAGGCAGCTGGGTGTGCTGGCCATGTCGGCCAGGAATGCCCGGTGGCCAG 357
QY 1094 GlySerGlyProIleThrLeuAspLeuAsnGlyThrGlyThrGluSerHisLeuTyr 1113
|||||
Db 358 GGCTCAGGACCATGTCCTGATGATGTGCTGCTGAGGAATGAGTCTACCTGTGG 417
QY 1114 GlnCysProSerArgGlyThrGlnHisAspCysArgHisLysGluAspAlaGlyVal 1133
|||||
Db 418 AGCTGCCCCCAAGAGCTGGCTCACCAATGTCGACATCAGGAGCAGCTGTGTC 477
QY 1134 IleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThr 1147
|||||
Db 478 ATCTGCTCA-----GCCACCAATAAATCTACTACGACAGATTGGTGGCATCCA 528
QY 1148 -----GluThrGluSerCysAlaGlyArgLeuGluValPheTyr 1160
|||||
Db 529 ACAACTACAACTACGACGCTCTTCAAAATGTTGGTGGCTCTTA-----TTCTAT 582
QY 1161 ---AsnGlyThrTrpGlySer 1166
|||||
Db 583 GCAGTGGGACATCTCCAGCCCATCTACCTGCTACATACCCCAACAATGCTAAGTGT 642
QY 1167 -----ValGlyArgArgAsnIle 1172
|||||
Db 643 GTTGGGAATAGAGTGAATCTGTTATCGCATAAACCTGGCTTCAGTAACTCGAAA 702
QY 1172 ----- 1172
Db 703 TTGGAGGCACACATAACTGCTGTTGATTATGTTGAATCTTGTGATGATCATTTGAAT 762
QY 1173 ThrThrAlaIleAlaGlyIleValCys-----ArgGluLeuGlyCysGlyGluAsn 1189
|||||
Db 763 ACAGTCTCTCTGGGAAATCTGTAATGATACAGCAATATTTACATCTCTTAC 822
QY 1190 GlyValValSerLeu-----AlaProLeuSerLysThrGlySerGlyPheMet 1205
|||||
Db 823 AACCGAATCACCATTACCTTTCGAAGTGACATCAGTTTCCAAACACAGCTGGCTTTGGCT 882
QY 1206 TrpValAspPheIleGlnCysProLysThrHisIleSerIleThrPheGlnCysLeuSerAla 1225
|||||
Db 883 TGGTATAAC----- 891
QY 1226 ProTrpGluArgArgIleSerProAlaGluGluThrTrpIleThrCysGluAspArg 1245
|||||
Db 892 -----TCTTCCCAACGATGCCACCTTG-----AGGTG 921
QY 1246 IleArgValArgGlyCysThrGluCysSerGlyArgValGluIleThrHisAlaGly 1265
|||||
Db 922 GTCAATTTAAATTCATCCCTATGCTATGTCGGGGGCTGAGAAATTTACCATGCTGTC 981
QY 1265 SerTrpGlyThrValCysAspSerTrpAspLeuAlaGluValValCysGln 1285
|||||
Db 982 ACCTGGGGACAGTTTGTGATGACTCTGACCATTTACGAAAGCTGAGTGTCTGCAGA 1041
QY 1286 GlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAlaSerPheGlyGlnGlyThr 1305
|||||
Db 1042 CAGTAGGCTGGAGCTGCAGCTTTCAGCCCTTGAATATGATATTTGGCTGCTGCTCT 1101
QY 1306 GlyThrIleTrpLeuAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCys 1325
|||||
Db 1102 GGCCCCATCACCTGGACCATGATGCTCTCAGGAGCGAATCCACTCTCTGGCAGTGC 1161

QY 1326 HisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCys 1345
|||||
Db 1162 CGGAACCGAGGTGTTCTCCCAACACTGAATCATCTGTAAGATGCTGTGTCATCTGC 1221
QY 1346 SerGlyGlnSerLeuLysSer 1352
|||||
Db 1222 TCAGGAACCATCTATCGACA 1242
RESULT 5
US-09-034-916-1
; Sequence 1, Application US/09034916
; Patent No. 6046314
; GENERAL INFORMATION:
; APPLICANT: GEBE, JOHN A.
; APPLICANT: SIADAK, ANTHONY W.
; APPLICANT: ARUFFO, ALEJANDRO A.
; TITLE OF INVENTION: SPALPHA: A NOVEL SCAVENGER RECEPTOR
; TITLE OF INVENTION: CYSTEINE-RICH DOMAIN-CONTAINING POLYPEPTIDE AND MONOCLONAL
; TITLE OF INVENTION: ANTIBODIES THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBINS & ASSOCIATES
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/034,916
; FILING DATE: 04-MAR-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,956
; FILING DATE: 06-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 5998-0020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 325-7812
; TELEFAX: (650) 325-7823
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2178 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 64..1101
US-09-034-916-1
Alignment Scores:
Pred. No.: 2,94e-54 Length: 2178
Score: 691.50 Matches: 148
Percent Similarity: 52.51% Conservative: 40
Best Local Similarity: 41.34% Mismatches: 139
Query Match: 8.50% Indels: 31
DB: 3 Gaps: 9
US-09-759-130B-381 (1-1453) x US-09-034-916-1 (1-2178)
QY 21 LeuPheSerAlaValValThrCysIleLeuLeuAsnSerCys-----Phe 36
|||||
Db 70 CTATCTCC-----TTGATCCCTTGCATTTGCACAGACCTGGATTC 111
QY 37 LeuIleSerSerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyPro 56

Db 112 CTAGCGTCTCCATCTGGA-----GTGCGCGTGGTGGGGGCTCCACCGC 156
Qy 57 CysSerGlyThrValGluValLysPheGlnGlyInrPrpGlyThrValCysAspGly 76
Db 157 TGTGAAGGCGGGTGGAGGTGGAAGAGAGCCAGTGGGACCGCTGTGTGATGACGGC 216
Qy 77 TrpAsnThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAla 96
Db 217 TGGACATTAAGGACGTGGCTGTGTGGCGGAGCTGGGCTGTGGAGTGGCCAGCGGA 276
Qy 97 Met-----PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAsp 113
Db 277 ACCCTAGTGTATTTGTATGATGACCCAGCAGCAAAAAGAGCAAAAGTCTCATCAA 336
Qy 114 AspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGly 133
Db 337 TCAGTCAAGTTCACAGGACAGAGATACATTGGCTCAGTGTGAGCAAGAA----- 390
Qy 134 SerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys----- 147
Db 391 GTTTATGATCTTCATGATGATGATGATGCTGGGCTGCTGGAGACCCAGAGCTCT 450
Qy 148 TyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArg 167
Db 451 TTCTCCCGAGTCCAGAGAGGTGTGAGGTGCTGAGGCTGAGGCTGAGGCTGAGGCT 510
Qy 168 ValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsn 187
Db 511 GTGGAAGTGAAGCAGCAGCAGTGGTATACCGTGTGCCAGAGAGGCTGGAGCTCCGG 570
Qy 188 ThrAlaAlaValCysArgGlnLeuGlyCysProSerSerPheIleSerSerGlyVal 207
Db 571 GCCGCAAGAGTGGTGTGCCGAGCTGGATGTGGAGGGCTGTACTGACTCAAAAGCC 630
Qy 208 ValAsnSerProAlaValLeuArg---ProIleTrpLeuAspAspIleLeuCysGlnGly 226
Db 631 TGCAACAAGCATGCTGTGCGGCGAAGCCATCTGGCTGAGCCAGATGCTATGCTCAGGA 690
Qy 227 AsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHis 246
Db 691 CGAGAGCAACCCCTTCAGGATGGCTTCTGGGCTTGGGGAGAGAACCTGCAACCAT 750
Qy 247 AsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuGluLeuArgLeuValGly 266
Db 751 GATGAAGACACGTGGTGTGATGTAAGATCCCTTTGAC-----TTGAGACTAGTAGGA 804
Qy 267 GlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyThrVal 286
Db 805 GGAGACAACCTCTGCTCTGGGCGACTGGAGGTGCTGCAACAAGGGCGTATGGGGCTCTGTC 864
Qy 287 CysHisHisLysTrpAsnAsnAlaAlaAspValValCysLysGlnLeuGlyCysGly 306
Db 865 TGTGATGACAACCTGGGAG 924
Qy 307 ThrAlaLeu-----HisPheAlaGlyLeuProHisLeuGlnSerGlySerAspValVal 324
Db 925 AAGTCCCTCTCTCCCTCTCCAGAGACCGGAAATGCTATGGCCCTGGGTTGGCCGATC 984
Qy 325 TrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCysArgHisSer 344
Db 985 TGGCTGGATAATGTTCGTTGCTCAGGGAGGAGGAGTCCCTGGAGCAGTGGCAGCAGACA 1044
Qy 345 GlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer 362
Db 1045 TTTTGGGGGTTTCACGACTGTCACCCAGGAGAGATGTGGCTGTCTGTCTCA 1098

RESULT 6

US-09-276-400-3

; Sequence 3, Application US/09276400

; Patent No. 6140056

; GENERAL INFORMATION:

; APPLICANT: Khodadoust, Mehran

; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
; FILE OF INVENTION: USES THEREFOR
; FILE REFERENCE: MNI-073
; CURRENT APPLICATION NUMBER: US/09/276.400
; CURRENT FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2262
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-276-400-3
Alignment Scores:
Pred. No.: 6,12e-49 Length: 2262
Score: 634.50 Matches: 214
Percent Similarity: 40.03% Conservative: 97
Best Local Similarity: 27.54% Mismatches: 253
Query Match: 7.80% Indels: 215
DB: 3 Gaps: 32

US-09-759-130B-381 (1-1453) x US-09-276-400-3 (1-2262)

Qy 30 LeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGlyThr----- 44
Db 40 CTGCTGCTG-----TGCCTGCTGTGCAGTTCGTGCTTGGGCTCTCCCTTCCACG 93
Qy 45 -----AspLeuGluLeuArgLeuValAsn---GlyAsp 54
Db 94 GCCCTGAGAGAGAGCGCGGAGCGGCTTCGGTCCCGCTGGCTGCCACAG 153
Qy 55 GlyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAsp 74
Db 154 AAGCCTACGAGGCGCGCTGGAGATACAGCAGCTGGTGAATGGGACCATCTGCGAT 213
Qy 75 AspGlyTrpAsnThrThrAlaSerThrValCysLysGlnLeuGlyCysProPheSer 94
Db 214 CATGACTTTCAGCTGTCAGGCTGCCACATCTCTCTCCGGAGCTGGGC-----TTCACA 267
Qy 95 PheAla-----MetPheArgPheGlyGlnAlaValThrArgHisGlyLys 109
Db 268 GAGGCCACAGGCTGACCCACAGCTGCCAATATGGCCCTGGGAACA-----GCCGC 318
Qy 110 IleTrpLeuAspAspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHis 129
Db 319 ATCTGGCTGGACAATCTGAGCTGCAGTGGGACCGAGAGAGTGTGACTGAATGCCCTCC 378
Qy 130 ArgGluTrpGlySerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys----- 147
Db 379 CGGGCTGGGGAGAACAGTGAAGTTCCTGACGCGCTGCTGCGGCTGCTATCTGCAAGAC 438
Qy 148 -----TyrGlyGluAlaAsnLeu----- 153
Db 439 CAGCCCTCCCTGGCTTCTCGGACTCCAATGCTAGGTAGAGCATCACCTGCAAGTG 498
Qy 154 -----GlyLeuArgLeuValAspGlyAsnAsn 162
Db 499 CAGGAGGTGCGAATTCGACCCCGCTGGTGGGCGGACAGACACCTCGCC----- 549
Qy 163 SerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAsp 182
Db 550 CTGACGGAGGGCTGGTGAAGTTCAGGCTTCCTGACGCGCTGCTGCGAAGTGTGCGACAA 609
Qy 183 GlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer---Ser 201
Db 610 GCGTGGAGCGCCACACAGCAGCAGCTGGTCTCGGGAGTGTCTCCCGGAGGAAAG 669
Qy 202 PheIleSerSerGlyValValAsn-----SerProAlaValLeuArgProIleTrp-Le 219
Db 670 AGGGTCAACGC---GGCTTCTACAGGCTGCTAGCCCCACCGCAGCAACATCTCTTGGTCT 728
Qy 219 uAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrp 239
Db 219 uAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrp 239

```

Qy 550 erAsnIleTrp---AspCysGluHisSerGlyTrpGlyHisAsnCysValHisArgG 569
|| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1799 CCTGGGTGGCCAGAGTGC-----CATGGCATTACCACAGCA 1837
Qy 569 luaspVal-----IleValThrCysSerGlyAspAlaThrTrpGlyLeuA 584
||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1838 TGGACATCTTCACACTCACTATGATATCTCACCCCAATGGCCACCAAGTG----- 1887
Qy 584 rgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTrpPheGlnGlyArgT 604
||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1888 -----GCTGAGGCGCCACA 1900
Qy 604 rpGlyThrValCysAspAspGlyTrpAsnSerIysAlaAlaValValCysSerGlnL 624
||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1901 AAGCTACTTTCTGT----- 1914
Qy 624 euaspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyL 644
||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1914 ----- 1914
Qy 644 ysIleTrpLeuAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgA 664
||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1915 -----CTCGAAGACACTGAGTGTGAGGAGGATGCTCTCCAAGCGGTATGAGTGTGCCA 1966
Qy 664 snSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerA 684
||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1967 ACTTTGGA-----GAGCAAGGCATCACTGTGGTGTCTGCTGGG 2020
Qy 684 sp-AlaSerAspMetGluLeuArgLeuValGlyGlySerSerArg 698
|| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 2003 ATCTCTACCGGCATGACATTGACTGTCACTGAGTGGATTGACATCAGG 2047

RESULT 7
US-09-448-076-3
; Sequence 3, Application US/09448076
; Patent No. 6300092
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran et al.
; TITLE OF INVENTION: METHODS OF USE OF A NOVEL LYSYL OXIDASE-RELATED PROTEIN
; FILE REFERENCE: WNI-073CP
; CURRENT APPLICATION NUMBER: US/09/448,076
; CURRENT FILING DATE: 1999-11-23
; EARLIER APPLICATION NUMBER: 60/117,580
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 09/276,400
; EARLIER FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2262
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-448-076-3

Alignment Scores:
Pred. No.: 6.12e-49 Length: 2262
Score: 634.50 Matches: 214
Percent Similarity: 40.03% Conservative: 97
Best Local Similarity: 27.54% Mismatches: 253
Query Match: 7.80% Indels: 215
DB: 4 Gaps: 32

US-09-759-130B-381 (1-1453) x US-09-448-076-3 (1-2262)
Qy 30 LeuLeuLeuAsnSerCysPheLeuIleSerPheAsnGlyThr----- 44
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 40 CTGCTGCTG-----TGCCTGCTGCTCAGTTCGTCTGGGGTCTCCGTCCTCCTTCCACG 93
Qy 45 -----AspLeuGluLeuArgLeuValAsn---GlyAsp 54
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 94 GGCCCTGAGAAGAAGCGCGGAGCCAGGCGCTTCGGTTCCGGCTCGGCTGCTTCCACAGG 153

```


Db 2003 ATCTACGGCATGACATGACTGCTCAGTGGATTGACATCACGG 2047

RESULT 8

US-09-702-572-3

; Sequence 3, Application US/09702572

; Patent No. 6391602

; GENERAL INFORMATION:

; APPLICANT: Khodadoust, Mehran

; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND

; TITLE OF INVENTION: USES THEREFOR

; FILE REFERENCE: MNI-073

; CURRENT APPLICATION NUMBER: US/09/702,572

; CURRENT FILING DATE: 2000-10-31

; PRIOR APPLICATION NUMBER: 09/276,400

; PRIOR FILING DATE: 1999-03-25

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 2262

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-702-572-3

Alignment Scores:

Pred. No.:	6,12e-49	Length:	2262
Score:	634.50	Matches:	214
Percent Similarity:	40.03%	Conservative:	97
Best Local Similarity:	27.54%	Mismatches:	253
Query Match:	7.80%	Indels:	215
DB:	4	Gaps:	32

US-09-759-130B-381 (1-1453) x US-09-702-572-3 (1-2262)

QY	30	LeuLeuAsnSerCysPheLeuLeuSerPheAsnGlyThr-----	44
Db	40	CTGCTGCTG-----TGCCTGCTGTCAGTTCGTGGGTCTCCGCTCCCTCCACG	93
QY	45	-----AspLeuLeuLeuArgLeuValAsn---GlyAsp	54
Db	94	GGCCCTTGAGAAGAGCCGGAGCCAGGGCTTCGGTTCCTGGCTGGCTTCGCCAGG	153
QY	55	GlyProCysSerGlyThrValGluValLysPheGlnGlyThrPglGlyThrValCysAsp	74
Db	154	AAGCCCTACAGGCGCGGTGAGATACAGCAGCTGGTGAATGGGGCCACATCTCGAT	213
QY	75	AspGlyTrpAsnThrAlaSerThrValCysLysGlnLeuGlyCysProPheSer	94
Db	214	GATGACTTCAGGCTGCAGCTGCCCATCTCTGCGGGAGCTGGGC-----TTCA	267
QY	95	PheAla-----MetPheArgPheGlyGlnAlaValThrArgHisGlyLys	109
Db	268	GAGGCCACAGCGCTGGACCCACAGTGCCTCAATATGGCCCTTGGAAACA-----	318
QY	110	IleTrpLeuAspValSerCysTyrglyAsnGluSerAlaLeuTrpGluCysGlnHis	129
Db	319	ATCTGCTGACAACTTGAGTGCATGAGCGGACGACAGAGTGAGTCAATGTGCTCC	378
QY	130	ArgGluTrpGlySerHisAsnCysTyrglyHisGlyGluAspValGlyValAsnCys	147
Db	379	CGGGGCTGGGGGAACAGTACTGTACCCAGCATGAGGATGCTGGGGTCTCTGCAAG	438
QY	148	-----TyrGlyGluAlaAsnLeu-----	153
Db	439	CAGCGCTCCCTGGCTTCCTGGACTCAATGTCATTGAGTAGAGCATCCTTCGAAGTG	498
QY	154	-----GlyLeuArgLeuValAspGlyAsnAsn	162
Db	499	GAGGAGTGGGAATTCGACCCCGCTGGGTGGGGGAGAGACCCCTGGCC-----	549
QY	163	SerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAsp	182
Db	550	GTACGAGGGGCTGGTGAAGTCAGGCTTCCTGACGGCTGGTTCGCAAGTGTGCGACAA	609

QY	183	GlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer-----	Ser 201
Db	610	GGCTGAGCGCCCAACACACACACGCTGCTGCGGATGCTGGCTCCCGCAAAAG	669
QY	202	PheIleSerSerGlyValValAsn-----SerProAlaValLeuArgProIleTrp-Le	219
Db	670	AGGTCAACGC--GGCTTCTACAGGCTGCTAGCCCAACGCGCAGCAACTCTTTGGTCT	728
QY	219	uAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTr	239
Db	729	GCATGGGTGGGTGGCGGCGGAGGCGCCACCTCTCCCTCTGTTCCCTGGAGTTCTA	788
QY	239	pGlyAsnHisAsp-----CysSerHisAsnGluAspValThrLeuThrCys-----	254
Db	789	TCGTGCCAATGACACCGCCAGGTGCCCTGGGGGGGCCCTGCAGTGGTGTGTGCC	848
QY	255	-----TyrAspSerSer-----	258
Db	849	AGGCCCTGTCTACGGCGCATCCAGTGGCCAGAAAGCAACACAGTCAAGCCTCAGGG	908
QY	259	-AspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLy	278
Db	909	GGAGCGCCCTGTGCTCTAAAGGGCGCGCCACCTGGAGAGGCGCGGTAGAAGTCT	968
QY	278	sIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAspVa	298
Db	969	GAAGCCACGACACATGGGCGACAGTCTGTACCGCAAGTGGGACCTGCATGCAGCGCT	1028
QY	298	lValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuG	318
Db	1029	GGTGTGTCGGGAGCTGGGCTTCGGGAGTCTCGAGAAGCTCTGAGTGGCGCTCGCATGG	1088
QY	318	nSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLe	338
Db	1089	CGAGGCGATGGGTGCTATCCACCTGAGTGAAGTTCGCTGCTGGCAGAGAGCTCTCCCT	1148
QY	338	uTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAspValSe	358
Db	1149	CTGGAAGTCCCGCCCAAGAACATCACAGTGAAGATTCTCACATGACGAGGATGCCGG	1208
QY	358	rValIleCysSer-----AspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySe	375
Db	1209	GTCTGGGTGCACTTACCTTACCTGGGCGAGAGACGATCCGACTCAGTGGGGCGG	1268
QY	375	rAsnAsnCysSerGlyArgValGluValArgIleHisGlu-----GlnTrpTr	391
Db	1269	CAGCCCAACATGAGGGCGAGTGCAGGTGCAATAGGGGACCTGGGCGCCCTTCGCTGGG	1328
QY	391	pThrIleCysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuG	411
Db	1329	CCTCATCTGTGGGATGACTGGGACCTGGAGGCCATGTTGGCTGTAGGCACTGGG	1388
QY	411	yCysProPheSerValPheGlySerArgAla-----LysProSerAsnGluAlaAr	429
Db	1389	TCTGGCTACGCCAACCCAGCGCTGCAGGAGACCTTGGTACTGGGACTCTGGGAATAA	1448
QY	429	gAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysTh	449
Db	1449	AGAGTGGTGTGATGGAGTGGCTGCACAGGAGTGAAGTGTCCCTGGATGAGTGTGC	1508
QY	449	rTyrrAspGlyLysAlaLysArgThrCysPheArg-----ArgSerAspAlaGlyVa	466
Db	1509	CATCATGCG-----ACCCATCCTCGAAGAGGAGGAGGAGCCCGCTTCACTGCTGGAGT	1565
QY	466	lIleCysSerAspLysAlaAspLeuArgLeuValGlyAlaHisSerProCysty	486
Db	1566	CATCTGTTCTGAGACTGCATCA---GATCTGTGCTG-----CACTCA-----	1605
QY	486	rglyArgLeuGluValLysTyrrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSe	506
Db	1606	-----GCATGGTGTGCA--GGAGCCGCGCTAC	1630
QY	506	rThrArgAsnAla-----AlaValValCysLysGlnLeuGlyCysGlyLysProMe	523


```

Db 1631 TCAGAACCGCCCGCTGCATATGTTGACTGTGC-----TGCAGAGAGAACT 1678
QY 523 t-----HisValPheGlyMetThrTyPheLysGluAlaSerGlyProIl 538
Db 1679 GCCTGGCCAGCTCAGCCCGCTCAGCCACCTATGCTATGCTCAGCGCTCTCCTCCGAT 1738
QY 538 eTrpLeuAspVal-SerCys-----IleGlyAsnGluS 550
Db 1739 TCTCTCCAGATCCACACACTGGGAGAGCTGACTTCAGGCCCAAGGCTGGGCCCACT 1798
QY 550 erAsnIleTrp---AspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArg 569
Db 1799 CCTGGGTGGCCAGCAGTGC-----CATGGGCATTACACACAGCA 1837
QY 569 luAspVal-----IleValThrCysSerGlyAspAlaThrTrpGlyLeuA 584
Db 1838 TGCATCTTCACTCACTATGATATCTCACCCCAATGGCACCAGGTG----- 1887
QY 584 rgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyPheGlnGlyArgT 604
Db 1888 -----GCTGAGGCCACCA 1900
QY 604 rpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValValCysSerGlnL 624
Db 1901 AAGCTAGTTCTGT----- 1914
QY 624 euAspCysProSerSerIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyGlyL 644
Db 1914 ----- 1914
QY 644 ysIleTrpLeuAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgA 664
Db 1915 -----CTCGAGACACTGAGTGTGAGGAGTGTCTCCAGCGGTATGAGTGCCCA 1966
QY 664 snSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerA 684
Db 1967 ACTTTGCA-----GAGCAAGGCATCACTGTGGTGTGCTGGG 2002
QY 684 sp-AlaSerAspMetGluLeuArgLeuValGlyGlySerSerArg 698
Db 2003 ATCTTACCGGCATGACATTGACTGTGAGTGTGATGACATCAGG 2047

```

RESULT 9

```

US-09-276-400-1
; Sequence 1, Application US/09276400
; Patent No. 6140056
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
; FILE OF INVENTION: USES THEREFOR
; FILE REFERENCE: MNI-073
; CURRENT APPLICATION NUMBER: US/09/276,400
; CURRENT FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2920
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (143)..(2401)
US-09-276-400-1

```

Alignment Scores:

```

Pred. No.: 9,29e-49 Length: 2920
Score: 634.50 Matches: 214
Percent Similarity: 40.03% Conservative: 97
Best Local Similarity: 27.54% Mismatches: 253
Query Match: 7.80% Indels: 215
DB: 3 Gaps: 32

```

```

US-09-759-130b-381 (1-1453) x US-09-276-400-1 (1-2920)
QY 30 LeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGlyThr----- 44
Db 182 CTGCTGCTG-----TGCCTGCTGTGAGTTCGTGTGGGTCTCGCTCCCTTCCACG 235
QY 45 -----AspLeuGluLeuArgLeuValAsn---GlyAsp 54
Db 236 GGCCTTGAGAAGAAGCGCGGAGCCAGGGGCTTCGGTTCGGCTGGCTTCCCGCAGG 295
QY 55 GlyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAsp 74
Db 296 AAGCCCTTACAGGCGCGCGTGAGATACAGCGAGCTGGTGAATGGGCGACCATCTCGAT 355
QY 75 AspGlyTrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysPropheSer 94
Db 356 GATGACTTACGCTGCGAGGCTGCCACATCTCTGCGCGGAGCTGGG-----TTCA 409
QY 95 PheAla-----MetPheArgPheGlyGlnAlaValThrArgHisGlyLys 109
Db 410 GAGGCCACAGGCTGGACCCACAGTGCCTCAATATGGCCCTGGAACA-----GGCCGC 460
QY 110 IleTrpLeuAspValSerCysTyPheGlyAsnGluSerAlaLeuTrpGluCysGlnHis 129
Db 461 ATCTGGCTGACAACTTGTGAGTGGAGTGGACCGAGCAGAGTGTGACTGAATGTGCTCC 520
QY 130 ArgGluTrpGlySerHisAsnCysTyPheHisGlyGluAspValGlyValAsnCys----- 147
Db 521 CGGGCTGGGGGAACAGTGTACGCACGATGAGGATGCTGGGGTCACTGCAAAAGAC 580
QY 148 -----TyrGlyGluAlaAsnLeu----- 153
Db 581 CAGGCCCTCCCTGGCTTCTCGGACTTCAATCTATTGAGGTAGAGCATCACCCTCAAGTG 640
QY 154 -----GlyLeuArgLeuValAspGlyAsnAsn 162
Db 641 GAGGAGGTGCGAATTCGACCCCGCGTGGTGGTGGGAGACAGCCCTCGCC----- 691
QY 163 SerCysSerGlyArgValGluValLysPheGlnGlyTrpGlyThrIleCysAspAsp 182
Db 692 GTGACGGAGGGCTGTGGAGTGTGAGTGTCTTCCGCGGTGGTGGCAAGTGTGGGACAAA 751
QY 183 GlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer---Ser 201
Db 752 GCTTGGAGCGCCCAACACAGCCAGCTGTCTGCGGATGCTGGGCTTCCCGCAGGCAAA 811
QY 202 PheIleSerSerGlyValValAsn-----SerProAlaValLeuArgProIleTrp-Le 219
Db 812 AGGTCAACGC-GGCCCTTCTACAGGCTGTAGCCCAACGGCAGCACTCCTTTTGGTCT 870
QY 219 uAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTr 239
Db 871 GCATGGGGTGGCTGGCGTGGCGACGAGGCCACCTCTCCCTCTGTCTCCCTGGAGTCTA 930
QY 239 pGlyAsnHisAsp-----CysSerHisAsnGluAspValThrLeuThrCys----- 254
Db 931 TCGTGCCAATGACACCGCCAGGTGCCCTGGGGGGGGCCCTGCAGTGGTGTGTGTGCC 990
QY 255 -----TyrAspSerSer----- 258
Db 991 AGGCCCTGTCTACGCGCATCTAGTGGCCAGAGAAGCAACACAGCTGAGAGCCTCAGGG 1050
QY 259 -AspLeuGluLeuArgLeuValGlyThrAsnArgCysMetGlyArgValGluLeuL 278
Db 1051 GAGGCCCGGTGTCTGCTAAAGGGCGGCGCCACCTCTGGAGAGCGCGGTAGAGTCT 1110
QY 278 sIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAspVa 298
Db 1111 GAAGGCCAGCACATGGGCGCAGTCTGTGACCGCAAGTGGGACCTGTGCTGAGCAGCGT 1170
QY 298 lValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuL 318
Db 1171 GTGTGTGCGGAGCTGGGCTTCGGGAGTGTCTCGAGAAAGTCTGTGAGTGGCGCTCGCATGG 1230

```

QY 318 nSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLe 338
Db 1231 GCAGGGCGATGGTGTATCCACCTGAGTGAAGTTCGCTGCTGCACAGGAGCTCCCT 1290
QY 338 uTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnValSe 358
Db 1291 CTGGAAGTCCCGCCCAAGAACAACATCACAGCTGAGGATGTTTCACATGACCGAGATGCCGG 1350
QY 358 rValIleCysSer-----AspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySe 375
Db 1351 GTCCTGGTGCACACCTTACCTTACCTGGGCGAGAGACAGGATCCGACTCAGTGGGGCGG 1410
QY 375 rAsnAsnCysSerGlyArgValGluValArgIleHisGlu-----GlnTrpTr 391
Db 1411 CAGCCCAACATGAGGGCGAGTGCAGGTGCANATAGGGGACCTGGGCCCTTCGCTGGGG 1470
QY 391 pThrIleCysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuG 411
Db 1471 CCTCATCTGTGGGGATGACTGGGGACCTGGAGCCATGTGTGGCTGTAGGCAACTGGG 1530
QY 411 yCysProPheSerValPheGlySerArgArgAla-----LysProSerAsnGluAlaAr 429
Db 1531 TCTGGCTAGCCACACCGCTGCAGAGACCTGCTGAGGACTCTGGGAATATATAC 1590
QY 429 gAspIleTrpLeuAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysTh 449
Db 1591 AGAGTGTGTGATGAGTGGAGTGCCTGCACAGGAGCTGCTCCCTGGATCAGTGTGC 1650
QY 449 rTrpAspGlyLysAlaLysArgThrCysPheArg-----ArgSerAspAlaGlyVa 466
Db 1651 CATCATGCTG-----ACCCACATACCTGCAAGAGGACAGGACCCCTTCACCTGTGGAGT 1707
QY 466 lIleCysSerAspLysAlaAspLeuAspLeuValGlyAlaHisSerProCysty 486
Db 1708 CATCTGTCTGAGACTGATCA--GATCTGTGTCTG-----CACTCA----- 1747
QY 486 rGlyArgLeuGluValLysTyrglnGlyTrpGlyThrValCysHisAspArgTrpSe 506
Db 1748 -----GCACTGGTGCA--GGAGACCGGCTACA 1772
QY 506 rThrArgAsnAla-----AlaValValCysLysGlnLeuGlyCysGlyLysProMe 523
Db 1773 TCGAAGACCGGCCCTGTATGTTGTACTGTG-----TGGCGAAGAGAACT 1820
QY 523 t-----HisValPheGlyMetThrTrpPheLysGluAlaSerGlyProI 538
Db 1821 GCTGGCCAGCTACCGCTCAGCACTGCGCTATGGTCACCGGCTGCTGCTCCGAT 1880
QY 538 eTrpLeuAspAspVal-SerCys-----IleGlyAsnGluS 550
Db 1881 TCTCCTCCAGATCCACAACTGGGAGAGCTGACTTCAGGCCCAAGGCTGGCGCCACT 1940
QY 550 eRAsnIleTrp---AspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgG 569
Db 1941 CCGTGTGGCAGCAGTGC-----CATGGGCAATACACAGCA 1979
QY 569 lAspVal-----IleValThrCysSerGlyAspAlaThrTrpGlyLeuA 584
Db 1980 TGGACATCTTACTACTATATATCCATCCCAATGGGACCAAGTG----- 2029
QY 584 rGLeuValGlyCysSerAsnArgCysSerGlyArgLeuGluValTrpPheGlnGlyArgT 604
Db 2030 -----GCTGAGGGCCACA 2042
QY 604 rPlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValValCysSerGlnL 624
Db 2043 AGCTAGTTTCTGT----- 2056
QY 624 euAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrglyL 644
Db 2056 ----- 2056

QY 644 ySileTrpLeuAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgA 664
Db 2057 -----CTCGAAGACACTGAGTGTCCAGGAGATGTCTCCAAGCGGTATGATGTGCCA 2108
QY 664 snSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSera 684
Db 2109 ACTTTGGA-----GAGCAAGGCATCACTGTGGTTGCTGGG 2144
QY 684 sp-AlaSerAspMetGluLeuArgLeuValGlyGlySerSerArg 698
Db 2145 ATCTCTACCGCATGACATGCTGCTCAGTGGATTCATCACGG 2189
RESULT 10
US-09-448-076-1
; Sequence 1, Application US/09448076
; Patent No. 6300092
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran et al.
; TITLE OF INVENTION: METHODS OF USE OF A NOVEL LYSYL OXIDASE-RELATED PROTEIN
; FILE REFERENCE: MNI-073CP
; CURRENT APPLICATION NUMBER: US/09/448,076
; CURRENT FILING DATE: 1999-11-23
; EARLIER APPLICATION NUMBER: 60/117,580
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 09/276,400
; EARLIER FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2920
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (143)..(2401)
US-09-448-076-1

Alignment Scores:
Pred. No.: 9,29e-49 Length: 2920
Score: 634.50 Matches: 214
Percent Similarity: 40.03% Conservative: 97
Best Local Similarity: 27.54% Mismatches: 253
Query Match: 7.80% Indels: 215
DB: 4 Gaps: 32

US-09-759-130B-381 (1-1453) x US-09-448-076-1 (1-2920)

QY 30 LeuLeuLeuAsnSerCysPheLeuLeuSerSerPheAsnGlyThr----- 44
Db 182 CTGCTGCTG-----TGCCTGCTGTGCAGTTCTGCTTTGGGCTCTCCGCTCCCTTCCACG 235
QY 45 -----AspLeuGluLeuArgLeuValAsn---GlyAsp 54
Db 236 GGCCCTGAGACAGAGCGCGGAGCGGCTTCGGCTCCGCTGGCTGGCTCCCCAGG 295
QY 55 GlyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAsp 74
Db 296 AAGCCTACGAGGCGCGGTGAGATACAGCGAGCTGCTGAATGGGACCATCTGCGAT 355
QY 75 AspGlyTrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSer 94
Db 356 GATGACTTACGCTGAGGCTGCCACATCTCTCTCCGGAGCTGGGC-----TTCACA 409
QY 95 PheAla-----MetPheArgPheGlyGlnAlaValIThrArgHisGlyLys 109
Db 410 GAGGCCACAGCTGGACCCACAGTGCACCAATATGGCCCTGGNAACA-----GGCCGC 460
QY 110 lIleTrpLeuAspValSerCysTyrglyAsnGluSerAlaLeuTrpGluCysGlnHis 129
Db 461 ATCTGGCTGGACAACCTTCACTGCTGAGTGGACCGACGAGAGTGTGACTGAATGTSCCTCC 520
QY 130 ArgGluTrpGlySerHisAsnCysTyrglyGluAspValGlyValAsnCys----- 147

Db 521 CGGGGCTGGGGAAACAGTGACTAGCCAGCATGAGGATGCTGGGTCTCATCTGCAAGAC 580
Qy 148 -----TyrGlyGluAlaAsnLeu----- 153
Db 581 CAGGGCTCCCTGGCTTCTCGGACTCCAATGTCTATTGAGTAGAGATACCTGCAAGTG 640
Qy 154 -----GlyLeuArgLeuValAspGlyAsnAsn 162
Db 641 GAGGAGTGGGAATTCGACCCCGCTGGGTGGGCGAGACGCCCTTGCCC----- 691
Qy 163 SerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAsp 182
Db 692 GTGACGGAGGGCTGGTGGGAAGTCAGGCTTCCTCAGCGCTGGTGGCAAGTGTGCGACAAA 751
Qy 183 GlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer---Ser 201
Db 752 GGTGGAGGCCCCACACACCGACGCTGGTCTGGGAGTGTGGGCTTCCCGCCAGGAAAAG 811
Qy 202 PheIleSerSerGlyValValAsn-----SerProAlaValLeuArgProIleTrp-Le 219
Db 812 AGGTCAACGC-GGCCTTCTACAGGCTGTAGCCCAACGCGCAGCAACACTCCTTTGTGCT 870
Qy 219 uAspAspIleLeuGlyCysGlnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrp 239
Db 871 GCATGGGGTGGCGTGGCGACGAGGAGCCACCTCTCCCTCTGTTCCTGGAGTTCTA 930
Qy 239 pGlyAsnHisAsp-----CysSerHisAsnGluAspValThrLeuThrCys----- 254
Db 931 TCGTGCCTCAATGACACCCCGAGTGCCTTGGGGGGGCGCCCTGCAGTGTGAGCTGTGTGCC 990
Qy 255 -----TyrAspSerSer----- 258
Db 991 AGGCCCTGTCTACGCGCATCCAGTGGCCAGAGAACAAACAGTCGAGCCTCAGGG 1050
Qy 259 -AspLeuLeuLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLys 278
Db 1051 GGAGGCCGTGTCCGTCTAAAGGCGCGCCACCCCTGGAGAGCGCGGTAGAAAGTCT 1110
Qy 278 sIleGlnGlyArgTrpGlyThrValCysHisLysTrpAsnAlaAlaAlaAspVa 298
Db 1111 GAAGCCAGCACATGGGGGCACGTCTGTGACCGCAAGTGGGACCTGCATGCAGCCACCGT 1170
Qy 298 lValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGl 318
Db 1171 GGTGTGCGGAGCTGGGCTTCGGGAGTGTGCGAGAGCTGTGAGTGGCGCTCGCATGG 1230
Qy 318 nSerGlySerAspValThrLeuAspGlyValSerCysSerGlyAsnGluSerPheLe 338
Db 1231 GCAGGGCATGGGTGCTATCCACCTGAGTGAAGTTCGCTGTGGACAGGAGCTCTCCCT 1290
Qy 338 uTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSe 358
Db 1291 CTGGAAGTGGCCCAAGACATCAGCTGAGGATGTTTCATAGCCAGGATGCGCG 1350
Qy 358 rValIleCysSer-----AspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySe 375
Db 1351 GGTCCGGTGCAACCTACCTTACAGTGGGCGAGAGACAGGATCCGACTCAGTGGGGCGG 1410
Qy 375 rAsnAsnCysSerGlyArgValGluValArgIleHisGlu-----GlnTrpTrp 391
Db 1411 CAGCCAACTAGGGGGGAGTCGAGGTGCAAAATAGGGGGAGCTGGGCGCCCTTCGCTGGG 1470
Qy 391 pThrIleCysAspGlnAsnTrpLysAsnGluAlaLeuValValCysLysGlnLeuGl 411
Db 1471 CCTCATCTGTGGGATCAGTGGGGGACCCCTGGAGGCCATGTGGCTGTAGGCAACTGGG 1530
Qy 411 yCysProPheSerValPheGlySerArgArgAla-----LysProSerAsnGluAlaAr 429
Db 1531 TCTGGGTACGCCCAACACCGCCCTCGAGAGACCTGTGTACTGGGACTCTGGGAATATAAC 1590
Qy 429 gAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysTh 449
Db 1591 AGAGGTGTGTAGTGTGAGTGCCTGACACAGGAGTGCAGTGTCCCTGGATCATGTGTC 1650

Qy 449 rTyrAspGlyLysAlaLysArgThrCysPheArg-----ArgSerAspAlaGlyVa 466
Db 1651 CCATCATGGC---ACCACATCCTGCAAGAGGACAGGACCGCTTCTACTGCTGGAGT 1707
Qy 466 lIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHisserProCysTy 486
Db 1708 CATCTGTCTGAGACTGCATCA---GATCTGTGCTG---CACTCA----- 1747
Qy 486 rGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSe 506
Db 1748 -----GCACGTGGTGA-GGAGACCCCTTACA 1772
Qy 506 rThrArgAsnAla-----AlaValValCysLysGlnLeuGlyCysGlyLysProSe 523
Db 1773 TCGAAGACCGGCCCTCATATGTTGACTGTGC-----TGCAGAGAGAAT 1820
Qy 523 t-----HisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIl 538
Db 1821 GCCTGGCCAGCTCAGCCCGCTCAGCCAACTGGCCCTATGTGTACCGCGTCTGCTCCGAT 1880
Qy 538 eTrpLeuAspAspVal-SerCys-----IleGlyAsnGluS 550
Db 1881 TCCTCTCCAGATCCACACACTGGGACGAGCTGAGCTCAGCCCAAGCTGGGCGCCACT 1940
Qy 550 eAsnIleTrp---AspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArg 569
Db 1941 CCTGGGTGTGGCAGAGTGC-----CATGGCATTACCACAGCA 1979
Qy 569 luAspVal-----IleValThrCysSerGlyAspAlaThrTrpGlyLeuA 584
Db 1980 TGGACATCTTCTACTCATATGATCTCTACCCCAATGGCACCAGGTG----- 2029
Qy 584 rGluValGlyGlySerAsnArgCysSerGlyArgLeuGluValTrpPheGlnGlyArgT 604
Db 2030 -----GCTGAGGCGCCACA 2042
Qy 604 rPglyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSerGlnL 624
Db 2043 AAGTACTGTTCTGT----- 2056
Qy 624 euAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyL 644
Db 2056 ----- 2056
Qy 644 ysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArg 664
Db 2057 -----CTCGAAGACACTGAGTGTGAGGAGGATGCTCCAGCGGTATGAGTGGCCA 2108
Qy 664 snSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerA 684
Db 2109 ACTTTGGA-----GAGCAAGGCATCACTGTGGTGTGCTGGG 2144
Qy 684 sp-AlaSerAspMetGluLeuArgLeuValGlyGlySerSerArg 698
Db 2145 ATCTCTACCGCATGACATTGACTGCTCAGTGGATGATGACACCG 2189

RESULT 11

US-09-702-572-1
; Sequence 1, Application US/09702572
; Patent No. 6391602
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: MNI-073
; CURRENT APPLICATION NUMBER: US/09702, 572
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 09/276, 400
; PRIOR FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1

```
; LENGTH: 2920
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (143)..(2401)
; US-09-702-572-1

Alignment Scores:
Pred. No.:          9,29e-49          Length:          2920
Score:             634.50             Matches:         214
Percent Similarity: 40.03%             Conservative:    97
Best local Similarity: 27.34%           Mismatches:     253
Query Match:       7.80%               Indels:         215
DB:                4                  Gaps:           32

US-09-759-130B-381 (1-1453) x US-09-702-572-1 (1-2920)

Qy 30 LeuLeuLeuAsnSerCysPheLeuLeuSerSerPheAsnGlyThr----- 44
Db 182 CTGCTGCTG-----TGCTGCTGTCAGTTGCTGCTTGGGCTCCCGTCCCGTTCCACG 235
Qy 45 -----AspLeuGluLeuArgLeuValAsn---GlyAsp 54
Db 236 GCCCTGAGAGAGCGCGGAGCGAGGGCTTCGGTCCGGCTGGCTGGCTTCCCGAGG 295
Qy 55 GlyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAsp 74
Db 296 AAGCCCTACGAGGCGCGGTGGAGATACAGCGAGCTGGTGAATGGGACCATCTCGCAT 355
Qy 75 AspGlyTrpAsnThrThrAlaSerThrValValCysLysGlnLeuLeuGlyCysProPheSer 94
Db 356 GATGACTTCACGCTCGAGGCTGCCACATCTCTCCGGGAGCTGGGC-----TTCACA 409
Qy 95 PheAla-----MetPheArgPheGlyGlnAlaValThrArgHisGlyLys 109
Db 410 GAGGCCACAGCTGACCCACAGTCACCAATGCCATATGCCCCCTGGACAC-----GGCCGC 460
Qy 110 IleTrpLeuAspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHis 129
Db 461 ATCTGGCTGGACAACATTGAGCTGCACTGGGACCGAGCAGAGTGCTGACTGAATGTGCCCTCC 520
Qy 130 ArgGluTrpGlySerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys----- 147
Db 521 CGGGGCTGGGGAGACAGTCACTGTCAGCAGATGAGGATGCTGGGGTCACTCTGCAAGAC 580
Qy 148 -----TyrGlyGluAlaAsnLeu----- 153
Db 581 CAGCGCTCCCTGGCTTCTCGACTCCCAATGTCATGAGGTAGAGCATCACCTGCAAGTG 640
Qy 154 -----GlyLeuArgLeuValAspGlyAsnAsn 162
Db 641 GAGGAGGTGCGAATTTCGACCCCGCTTGGGTGGGCGACAGCACCCCTGCC----- 691
Qy 163 SerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAsp 182
Db 692 GTGACGGAGGGCTGGTGAAGTACAGGCTTCTGACGGCTGGTCCCAAGTGTGGCACA 751
Qy 183 GlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuLeuGlyCysProSer---Ser 201
Db 752 GCGTGGAGCGCCCAACACAGCCACCTGCTGCGGGATGCTGGGCTTCCCGACGCAAAAG 811
Qy 202 PheIleSerSerGlyValValAsn-----SerProAlaValLeuArgProIleTrp-Le 219
Db 812 AGGGTCAACGC--GGCTTTCTACAGCTGCTAGCCCAACAGCGGACGACACTCTTTGGTCT 870
Qy 219 uAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTr 239
Db 871 GCATGGGTGGGTGGCGGACCGGAGCCACCTCTCCCTCTCTTCCCTGGAGTTCTA 930
Qy 239 pgLyAsnHisAsp-----CysSerHisAsnGluAspValThrLeuThrCys----- 254
Db 931 TCGTGCCCAATGACACCGCCAGGTGGCTTGGGGGGGGCCCTGCGAGTGGTGGCTGTGTGCC 990
```

```
Qy 255 -----TyrAspSerSer----- 258
Db 991 AGGCCCTGTCTACGGCGCATCCAGTGGCGAGAGAAAGCAACACAGTGCAGAGCCTCAGGG 1050
Qy 259 -AspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLy 278
Db 1051 GGAGGCCCGTGTCCGCTCTAAAGGGCGCGCCACCCTCGAGAGGCGCGGTAGAAGTCTCT 1110
Qy 278 stIeGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAspVa 298
Db 1111 GAAGCCACGACATGGGCGACAGTCTGTGACCGCAAGTGGGACCTGCATGCAGCCAGCGT 1170
Qy 298 lValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGl 318
Db 1171 GGTGTGTGCGGAGCTGGCTTCGGGAGTCTCGAGAAGCTCTGAGTGGCGCTCGCATGG 1230
Qy 318 nSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLe 338
Db 1231 CGAGGGCATGGGTGCTATCCACCTGAGTCAAGTTCGCTCTGCACAGGAGCTCTCCCT 1290
Qy 338 uTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSe 358
Db 1291 CTGGAAGTGGCCCCACAGAACATCACAGCTGAGGATGTTTCACATAGCCAGGATGCCGG 1350
Qy 358 rValIleCysSer-----AspGlyAlaAspLeuGluLeuArgLeuAlaAspCys 375
Db 1351 GGTCCGGTGCACCTTACCTTACACTGGGCGAGACAGGATCCGACTCAGTGGGGCGG 1410
Qy 375 rAsnAsnCysSerGlyArgValGluValArgIleHisGlu-----GlnTrpTr 391
Db 1411 CAGCCAAATGAGGGCGAGTGGAGGTGCAATAGGGGACCTGGGGACCTGGGCCCTTCCGTGGG 1470
Qy 391 pThrIleCysAspGlnAsnTrpLysAsnGlnAlaLeuValValCysLysGlnLeuGl 411
Db 1471 CCTCATCTGTGGGATGACTGGGGACCTGGAGGCCATGGTGGCTGTAGGCAACTGG 1530
Qy 411 yCysProPheSerValPheGlySerArgAla-----LysProSerAsnGluAlaAr 429
Db 1531 TCTGGCTACGCCAACACACCGCTGCAGGAGACCTGGTACTGGGACTCTGGGAATATAAC 1590
Qy 429 gAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysTh 449
Db 1591 AGAGTGTGTGAGTGGAGTGGCTGCACAGGAGTCTGAGTGTCCCTGGATGAGTGTGC 1650
Qy 449 rTyrAspGlyLysAlaLysArgThrCysPheArg-----ArgSerAspAlaGlyVa 466
Db 1651 CCATCATGCG---ACCCACATCACCTGCAAGAGGACAGGACCGCCCTTCACTGCTGGAGT 1707
Qy 466 lIleCysSerAspLysAlaAspLeuArgLeuValGlyAlaHisSerProCystTy 486
Db 1708 CATCTGTCTTGAGACTGCATCA---GATCTGTGTGTG---CACTCA----- 1747
Qy 486 rGlyArgLeuGluValLysTyrGlnGlyLutrpGlyThrValCysHisAspArgTrpSe 506
Db 1748 -----GCATGGTGTGCA--GGAGCCGCGCTACA 1772
Qy 506 rThrArgAsnAla-----AlaValValCysLysGlnLeuGlyCysGlyLysProMe 523
Db 1773 TCGAAGACCGGCCCTCGCATATGTTGTACTGTGC-----TGGGGAAGAGAACT 1820
Qy 523 t-----HisValPheGlyMetThrThrPheLysGluAlaSerGlyProIle 538
Db 1821 GCCTGGCCAGCTACGCCGCTCAGCAACTGGCCCTATGGTCACCGGCTGTGCTCCGAT 1880
Qy 538 eTrpLeuAspAspVal-SerCys-----lIleGlyAsnGlu 550
Db 1881 TCTCTCCAGATCCACAACCTGGGACGAGCTGACTTACGGCCCAAGGCTGGGGCCACT 1940
Qy 550 eRAsnIleTrp---AspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgG 569
Db 1941 CTGGGTGTGGCACGAGTGC-----CATGGGCATTACCACAGCA 1979
```

```
Qy 569 luAspVal-----lleValThrCysSerGlyAspAlaThrTrpGlyLeuA 584
Db 1980 TGGACATCTTCACTACATATATCTCACCACCAATGGACCAAGGTG----- 2029
Qy 584 rgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyPheGlnGlyArgT 604
Db 2030 -----GCTCAGGGCCACCA 2042
Qy 604 rpGlyThrValCysAspGlyTrpAsnSerLysAlaAlaValValCysSerGlnL 624
Db 2043 AAGTAGTTCTGT----- 2056
Qy 624 euAspCysProSerSerIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrglyL 644
Db 2056 ----- 2056
Qy 644 ysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArg 664
Db 2057 -----CTCGAAGACACTGAGTCTCAGGAGGATGCTCCACCGGTATGAGTGCCTCA 2108
Qy 664 snSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerA 684
Db 2109 ACTTTGA-----GAGCAAGGCATCAGTGTGGTGTCTGGG 2144
Qy 684 sp-AlaSerAspMetGluLeuArgLeuValGlyGlySerSerArg 698
Db 2145 ATCTCTACCGGCATGACATTGACTGTGCTCAGTGTGATGATGATCAGCG 2189
RESULT 12
US-09-341-587-7
; Sequence 7, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341,587
; EARLIER FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: PCT/DE98/00096
; EARLIER FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 28720
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-341-587-7
Alignment Scores:
Pred. No.: 1.18e-45 Length: 28720
Score: 618.50 Matches: 386
Percent Similarity: 33.63% Conservative: 183
Best Local Similarity: 22.81% Mismatches: 501
Query Match: 7.60% Indels: 629
DB: 4 Gaps: 80
US-09-759-130B-381 (1-1453) x US-09-341-587-7 (1-28720)
Qy 25 ValValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGly---- 43
Db 2872 ATTTGTTGACCTCTCTGGGGGATGATGAAGGTCTTTGTTCTCCCTGTAGGATCTGAA 2931
Qy 44 ThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThrValGluVal 63
Db 2932 TCCACTTTGGCCCTGAGACTGGTGAATGGAGGTGACAGGTGTCGAGGCCGAGTGGAGTGC 2991
Qy 64 LysPheGlnGlyIntrpGlyThrValCysAspGlyTrpAsnThrThrAlaSerThr 83
Db 2992 CTATACCAAGGCTCTCTGGGGCCCGCTGTGTGATGACTACTGGACACCAATGATGCCAAC 3051
Qy 84 ValValCysLysGlnLeuGlyCysProPheSerPheAlaMet-----PheArgPhe 100
Db 3052 GTGGTCTGACGACGCTGGGCTGTGGTGGGCCATGTTCAGGCCCCAGGAAATGCCAGTTT 3111
```

```
Qy 101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrglyAsn 120
Db 3112 GCCCAGGCG-----TCAGGACCCATCTCTCGATGATGCGCTGCTCAGGACAC 3162
Qy 121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrglyHisGly 140
Db 3163 GAGTCTTACCTGTGGAGTGCCTCCACCAATGGCTGCTGCCACAACTTGGCCATCAT 3222
Qy 141 GluAspValGlyValAsnCysTyrgly-----GluAlaAsnLeuGly 154
Db 3223 GAAGATGCTGTGTGTCATCTGCTCAGGTGGGCTTTCAAGACCTTGGGCTCCCTCTCTTAAG 3282
Qy 155 LeuArgLeuValAspGlyAsnAsn----- 162
Db 3283 TTGAAGTTGCTCAGGAAGAAATCCTAATTACATTCTGATCTCTCACTCAAAAGCTTTT 3342
Qy 163 -----SerCysSerGlyArgValGluValLysPheGlnGlu 174
Db 3343 TCTATGTTTTCTATATTTCTGAAGTCTTGTAGCTCTCTGCTAAGAAATCTTTATGAATTT 3402
Qy 175 ArgTrpGlyThrIleCysAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArg 194
Db 3403 TGCTACAGTACTCTGGTGCAGC-----TGTGGC 3429
Qy 195 GlnLeuGlyCysProSerSerPheIleSerSerGlyValValAsnSerProAla----- 212
Db 3430 CACTTAGCCAGGCTCCGAACCTGAA-ACAACAACCCAGACTTATCCCTCCCTCCTGAGGC 3488
Qy 213 -----ValLeuArgProIleTrpLeuAspIleLeu 223
Db 3489 AGTGAAGGAAGAGGACAGAGAGAAAGTCTGGCTCC----- 3527
Qy 224 CysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAsp 243
Db 3528 ---CACGGCTCCATTTCTTCCCTGCTGAGTAGCAGTGGTTGAGGGTATCTGGGACACAGC 3584
Qy 244 -----CysSerHisAsnGluAspVal---ThrLeuThrCys 254
Db 3585 ACAGATAGCAGGGGCGAGGGAGGATCTCTCCTACACAGGAACCTGTGAACATAA-AGATGC 3643
Qy 255 TyrAspSerSer-----AspLeuGluLeuArg 263
Db 3644 TTGCTCGAAGTGGTCTCAGCTCAGACCCAGTGCAGAGGTCTGGAATAGAGCTCAA 3703
Qy 264 LeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGlyArgTrp 283
Db 3704 GGGTTAGGAGT-GCAATGTCATGT-----CTGTTGTCTCAGGCTGCTT 3747
Qy 284 GlyThrValCysHisHisLysTrpAsnAsnAlaAlaAspValValCysLysGlnLeu 303
Db 3748 GGAGTCTCTGACCTCAGGCTCCTTAAGAATGCTGCAGAGCAGTGCCTGCTGCC-CAGGTC 3806
Qy 304 GlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGlySerAspVal 323
Db 3807 TGGTGTGGG-----GAGGCGACGCCCAT---GAGTCTGGCCAG----- 3842
Qy 324 ValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCys----- 341
Db 3843 -----GCATGGCTGTGTTATTGC 3860
Qy 342 -----ArgHis----- 343
Db 3861 CTGTGGTGGGGCTGTAGATCACAAGAGCATTTGGGGTGGAGTGGCTCTCCTCAGCCTT 3920
Qy 344 -----SerGlyThr-----ValAsnPheAsp 350
Db 3921 GCTGACTCAGGAACCTGCCAAAACATCCAAAGGAGAGTGTGTTTGGGTCAACCTGGTC 3980
Qy 351 CysLeuHisGlnAsnAspValSerValIleCysSerAsp-----GlyAlaAsnLeuGlu 368
Db 3981 ACCCTGGCAGACACAAAGTTACTCACCTCGAGCTGACAATAGTGGGCGAGATCTGCCT 4040
```

Qy	369	LeuArgLeuAlaaspGlySerAsnAsnCysSerGlyArgValGluValArgIleHisGlu	398
Db	4041	GCACCCCTTATA-----TGGTCGATCTCTGGGAATTTACATG	4079
Qy	389	-GluTrpTrpThrIleCysAspGln-----AsnTrpLysAsnGluGlnAl	403
Db	4080	GCAATGCCCTCCTCTGTATAGGAGCTAGGACTGAGTGTGAGGCTTGCCAGTT	4139
Qy	403	aLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaL	423
Db	4140	CCTTCTATCTTTGTCACGTTT---TGCCATTTTCTGTAT-----	4176
Qy	423	sProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSe	443
Db	4177	-----AGTGCATCTGACTGACCTTCTCTCTCTTCT-----	4209
Qy	443	rAlaLeuTrpAspCysThrTyrrAspGlyLysAlaLysArgThrCysPheArgArgSerAs	463
Db	4210	-----CACAGCTGCTCAGTCCAGTCCAGTCAACGCCAG-----GCCAGGTGAGT	4250
Qy	463	pAlaGlyValIleCysSerAspLysAlaAspLeuArgLeuValGlyAlaHisSe	483
Db	4251	CCCCAGCATCTTCATCGGAGTGCCTTCTCTTCTGCCAGGTACCTCTTCCCCACTC	4310
Qy	483	rProCysTyrrGlyArgLeuGluValLys-TyrrGlnGly-----	495
Db	4311	CACAGAGCTCTCCTGCTTTTCTGTGGGATACTGTTGGGCATATATATTTTCTCCACC	4370
Qy	496	-----GluTrpGlyThrValCysHisAspArTrpSerThrArgAsnAlaVal	512
Db	4371	ACTCTGTAACGTAGAGCCCGACGATATGCTTCAACAGACATAGAGGTTCAGGAGC	4430
Qy	512	alValCys-LysGlnLeu---GlyCysGlyLysProMetHisValPheGlyMetThrTy	530
Db	4431	TCCTCTGTTCAATTTATTTTCAGGTTTCATGAAGCAGTTTCATACTGTACATGGCA	4490
Qy	531	PhelysGlnAlaSerGlyProIleTrpLeuAspAspValSerCysIleGlyAsnGluSer	550
Db	4491	CTTACAGGTTTCAGGAAGTGGCCTCATGTTTAAATAGCTTTAGCTCAATTTATTCAG	4550
Qy	551	AsnIle-----TrpAspCysGluHis	557
Db	4551	TG-ATACAACCTTTCGAGAAITGAGCTGACTGCCAAGTCACTGGGAGGCAATGTC	4609
Qy	558	SerGlyTrpGlyLysHisAsnCysValHisArgGluAspValIleVal-----	573
Db	4610	AGTTCAAGCCTTAAACATGACTCTGCCATGGCAAGCAATTCAGTGCAGCGCTGATAC	4669
Qy	573	-----	573
Db	4670	CTCGTCCCTCACTCCTTCAATACCCAGATTTTGCAGGCCATATCTGCATCCAGAAA	4729
Qy	574	-----ThrCysSerGlyAspAlaThr-----TrpGlyLeuArgLeu	585
Db	4730	AGGCAGAGCCATGCTCGGCGAGGAGGGGATAATAAATATTTCTGTGTGCTCCACTTA	4789
Qy	586	valGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrrPheGlnGlyArgTrp---	604
Db	4790	-----TCAGAAACTTTGATACCCCTTTGGGGGCTTCCTTGTGTTTC	4828
Qy	605	-----GlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValVal	620
Db	4829	CCCTAACATTTTAGTCGCACTGT-----CAGAGTCTCAGCAATGGTGTCCACA	4876
Qy	621	CysSerGlnLeuAsp-CysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerTh	640
Db	4877	TGTGCCCATCAGTCCATGT-----	4900
Qy	640	rGlyTyrrGlyLysIleTrpLeu-----AspAspValSerCysAspGlyAspG	656
Db	4901	TGGATGGGCGAGGCTACCCCTTCGTAGTCGAGAAGAGCATCTCACTTACAGAGTAG	4960
Qy	656	userAspLeuTrpSerCysArgAsnSer-----GlyTrpGlyAsnAsnAspCysSerHi	674

Db 5738 GTCACAGG-----TGCTTCCCAATACTTGGAGCTTCCATAGACTTGGGTGGAGTAGGAC 5791
Qy 978 --AsnCys-GlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSer 996
Db 5792 ATCACTGCTTCTCCATATCATCAAGCTGACCTCTGTCTGTATTCATATTCATCAAGGTG 5851
Qy 997 ValIleCysThr-----GlySerLeuThrGln-Pro----- 1006
Db 5852 ATTACCTGCACACGCTGACTTTGGCCCAATTAGGAAGTGCCTGTAGTGGAAACATTCCTTTA 5911
Qy 1007 -LeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaValProGluGlySe 1026
Db 5912 AATCCTTGACCTCATATCATGATGGATGAAGGCTTCTGTGTCTCC---CCTGTAGGATC 5968
Qy 1026 rAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAspSerArgCysAl 1046
Db 5969 TGAA-----TCCAGTTTGGCTCTGAGGCTGGTGAATGGAGGTGACAGGTGCTG 6016
Qy 1046 aGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAs 1066
Db 6017 AGGCCGAGTGGAGGTCTCTGTATCGAGGCTCTGGGGAACCGTGTGTATGACAGCTGGGA 6076
Qy 1066 pLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaTh 1086
Db 6077 CACCAATGATGCCAATGGTGTGACGACAGCTGGCTGTGGCCATGTGCGGCC 6136
Qy 1086 rValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspLeuAsnCysTh 1106
Db 6137 AGGAATGCCGGTTTGGCCAGGCTCAGGACCCATTTGCTGTGATGATGTGCGCTGCTC 6196
Qy 1106 rGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysAr 1126
Db 6197 AGGGAATGAGTCTACCTGTGGAGCTGCCCCACAAAGGCTGGCTCACCCACAACTGGG 6256
Qy 1126 gHisGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGl 1146
Db 6257 CCATCAGCAAGACGCTGGTGTCTCTGCTCA----- 6287
Qy 1146 uThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySe 1166
Db 6288 -----GGTGGGCC--TTCAGACCTGGGCTCCCTCTCTTGGGGTGG 6327
Qy 1166 rVal---GlyArgArgAsn-----lleThrAlaIleAlaGl 1178
Db 6328 AGTTTGCTCCAGAGAAACTCTTAATTCTGATCTCTCTCACTCAAGCTTCTCTCGT 6387
Qy 1178 yIleValCys----- 1181
Db 6388 TGTTTCTCTGCTTTTGAAGACTTGTAGCTCTCTGTAAGATCCATATGAATTCACGT 6447
Qy 1182 -----ArgGlnLeuGlyCysGlyGluAsnGlyValValSer----- 1193
Db 6448 CCTAGTGTCTCTGGTCTACTTAGCAGGGGACCAACTCAACACACCCAGACTTT 6507
Qy 1194 -----LeuAlaPro----- 1196
Db 6508 ATCCCTTCTCTGAGCACTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6567
Qy 1197 -----LeuSerLysThrGlySerGlyPheMetTrpValAspIleGl 1211
Db 6568 CTCCATTTCTCCCTACTGAGTAGACTGGGTGAGGTATGG-TGGACACAGCAGACAGA 6626
Qy 1211 nCysProLysThrHisIle-----SerIleTrpGlnCysLeuSerAlaPr 1226
Db 6627 CGGGGAGCAGAGGAGGATCTCTCTCTAGGAGGAGGCTCATGGTAAGGAAGGACATATG 6686
Qy 1226 oTrpGluArgGlyIleSer-----SerProAlaGluGluThrTrpIleTrpCysGluAs 1244
Db 6687 TTGGGTAGGAATCTCTACTTCTAGGCCCCAGTAGGAGGACATTTGGATT----- 6735
Qy 1244 pArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAl 1264
Db 6736 -----GGAGGC-----ATATGGGCTAA 6752

Qy 1264 aGlySerTrpGlyThrValCysAspSerTrpAspLeuAlaGluAlaGluValValCy 1284
Db 6753 GAGTGCAACGGGTATCTGTGC-----ATGTG 6779
Qy 1284 sGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGlyGlnGl 1304
Db 6780 TGAC-----CTGGG 6788
Qy 1304 yThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsn-----GluSerPh 1321
Db 6789 TCTTGGTCAAGTTTCAGCTCTCCCTCCAAAGAACTGCAGTGCATGTGTCTAGCAAGCATG 6848
Qy 1321 eLeuTrpAsp-----CysHisAlaLysPro-----Tr 1330
Db 6849 GTGTTGGGAGGCGAGCTCCCATGAGATCTGCAGGCAAGGCTTGTATTACGTGTGTT 6908
Qy 1330 pGlyGlnSerAspCysGlyHisLysGluAsp-----AlaGly----- 1342
Db 6909 TGGGATGGAAGT-----CACACAGGGGATTTTGTCTGGAGTGTATCTCTGAGACTTGC 6962
Qy 1343 -----ValArgCys-----SerGlyGlnSerLeuLysSerLeuAsnAl 1355
Db 6963 TGACTTGTGTAATTGCTAAACCTGCAAGGGAAGGTTGTTTGTTCATTGACATCTGCTGG 7022
Qy 1355 aSerSerGlyHis----- 1361
Db 7023 CACATCAGACACAGAGTTAATTCCTTGGAGCTGGAATGTTGGACAGAACTGCTGTTGG 7082
Qy 1361 aLeuIleLeuSerSerIlePhe 1368
Db 7083 ATCCCTTATAAGGAGCGCTTTT 7104
RESULT 13
US-08-477-674-9
; Sequence 9, Application US/08477674
; Patent No. 5644035
; GENERAL INFORMATION:
; APPLICANT: Koths, Kirston E.
; APPLICANT: Halenbeck, Robert E.
; APPLICANT: Taylor, Eric W.
; APPLICANT: Wang, Alice M.
; APPLICANT: Caspitt, Clayton L.
; TITLE OF INVENTION: Secreted Mac-2-Binding Glycoprotein
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cetus Oncology Corporation
; STREET: 1400 Fifty-Third Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,674
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/316,714
; FILING DATE:
; APPLICATION NUMBER: US/07/961,404
; FILING DATE: 15-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Kenneth M.
; REGISTRATION NUMBER: 34,174
; REFERENCE/DOCKET NUMBER: 2595.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 420-3152
; TELEFAX: (510) 658-5470

Best Local Similarity: 30.57% Mismatches: 108
Query Match: 3.86% Indels: 96
DB: 1 Gaps: 15

US-09-759-130B-381 (1-1453) x US-08-473-791-9 (1-2285)

```
QY 1031 LeuGluAspLysArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGlu 1050
DB 234 GTGAATGATGTGACATGCGCTGGCGATGGGGCGCCACCAACAGGCGCGGTGGAG 293
QY 1051 IleTyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAla 1070
DB 294 ATCTTCTACAGAGCCAGTGGGCGACTGTGTGTGACAACTGTGGGACCTGACTGATGCC 353
QY 1071 HisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHis 1090
DB 354 AGCGTCTGTCGGGGCGCTGGGCTTCGAGACGCCACCCAGGCTCTGGCGAGAGCTGCC 413
QY 1091 PheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSer 1110
DB 414 TTCGGCAAGGATCAGGCCCATCATCTGCGACGAGGTCCAGTGCACGGGAACCGAGGCC 473
QY 1111 HisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAsp 1130
DB 474 TCAGTGGCCGACTGCAAGTCCCTGGGCTGGCTGAAGAGCAACTGCAGCGACGAGAGAGC 533
QY 1131 AlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGlu 1150
DB 534 GCTGGTGTGCTGCACCAATGAACC-----AGGACACCCACACCCCTG 578
QY 1151 SerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgArg 1170
DB 579 GACCTCTCC-----AGG 590
QY 1171 AsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGly 1190
DB 591 GAGCTCTGGAGGCCCTTGGCCAGATCTTGCACAGCCAGGGGCTCGACCTGTCC--- 647
QY 1191 ValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspIle 1210
DB 648 ---ATCAGCGTGAATGTGCAGGGGAGGAGCCCTGGGCTC----- 686
QY 1211 GlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArg 1230
DB 687 ---TGTGGCCACACGGTCATC-----CTGACTGCCAACCTGGAC----- 722
QY 1231 IleSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArgValArgGly 1250
DB 723 -----GCCAGGCCCTGTGAAGAGCGCGGCAATGTCACCATG---AGT 767
QY 1251 GlyAspThrGluCys-----SerGlyArg 1258
DB 768 GTGGATGCTGAGTGTGTGCCATGCTCAGGAGCCTTCTCAGGTACTTCTACTCCCGAAG 827
QY 1259 ValGluIle-----TrpHisAlaGlySerTrpGlyThrValCys 1271
DB 828 ATTGACATCACCTCTGCTGCTCAGTCAAGTGTCTCCAC----- 863
QY 1272 AspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGly---CysGly 1290
DB 864 -----AGCTGGGCTCTGCTATGGGGCCAGGAGCGTCGAGGGCTACTGCGCA 911
QY 1291 SerAlaLeuAlaLeu-----ArgAspAlaSerPheGlyGlnGlyThrGlyThrIle 1308
DB 912 AGCCTTTGGCATCTCTCTCCCGCAGGACCCCTCGTTCCA-GATGCCCTGGACCTGTA 970
QY 1309 TrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLys 1328
DB 971 TGCCTA-----TGACGTGGCCAC-----AGGGGACGC 997
QY 1329 ProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGln 1348
DB 998 CCTGCTGGA-----GAAGCTCTGCTACAGTTCCTGGCCTG 1033
```

QY 1349 SerLeuLysSerLeuAsnAlaSerSerGly 1358
DB 1034 GAACCTTCGAGGCTTTGACGACGAGCGGAGGC 1063

RESULT 15

```
US-08-316-714-9
; Sequence 9, Application US/08316714
; Patent No. 5965382
; GENERAL INFORMATION:
; APPLICANT: Koths, Kirston E.
; APPLICANT: Halenbeck, Robert F.
; APPLICANT: Taylor, Eric W.
; APPLICANT: Wang, Alice M.
; APPLICANT: Caspitt, Clayton L.
; TITLE OF INVENTION: Secreted Mac-2-Binding Glycoprotein
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cetus Oncology Corporation
; STREET: 1400 Fifty-Third Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/316,714
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/961,404
; FILING DATE: 15-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Kenneth M.
; REGISTRATION NUMBER: 34,174
; REFERENCE/DOCKET NUMBER: 2595.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 420-3152
; TELEFAX: (510) 658-5470
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2285 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-316-714-9
```

Alignment Scores:
Pred. No.: 3,56e-19 Length: 2285
Score: 314.00 Matches: 107
Percent Similarity: 42.00% Conservative: 40
Best Local Similarity: 30.57% Mismatches: 108
Query Match: 3.86% Indels: 96
DB: 2 Gaps: 15

US-09-759-130B-381 (1-1453) x US-08-316-714-9 (1-2285)

```
QY 1031 LeuGluAspLysArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGlu 1050
DB 234 GTGAATGATGTGACATGCGCTGGCGATGGGGCGCCACCAACAGGCGCGGTGGAG 293
QY 1051 IleTyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAla 1070
DB 294 ATCTTCTACAGAGCCAGTGGGCGACTGTGTGTGACAACTGTGGGACCTGACTGATGCC 353
QY 1071 HisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHis 1090
```

Db 354 AGCGTGTCTGCGGGCCCTGGGCTTCGAGAACGCCACCGAGTCTGGGAGAGCTGCC 413
QY 1091 PheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSer 1110
Db 414 TTCGGGCAAGGATCAGGCCCATCATGCTGGACGAGTCCAGTCCACGGGACCGAGGCC 473
QY 1111 HisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAsp 1130
Db 474 TCACGTGGCGGACTGCAAGTCCCTGGCTGGCTGAAGACCACTGACGAGCAGCAGAGAGAC 533
QY 1131 AlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGlu 1150
Db 534 GCTGTGTGTGCTGTCACCAATGAAACC-----AGGAGCACCACACACCTGTG 578
QY 1151 SerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgArg 1170
Db 579 GACCTCTCC-----AGG 590
QY 1171 AsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGly 1190
Db 591 GAGCTCTCGAGGCCCTTGGCCAGATCTTTGACAGCCAGCGGGGCTGGCACCCTGTCC--- 647
QY 1191 ValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAspIle 1210
Db 648 ---ATCAGCGTGAATGTGCGGGCGAGGAGCGCCCTGGGCTTC----- 686
QY 1211 GlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArg 1230
Db 687 ---TGTGGCCACACGGTCATC-----CTGACTGCCAACCTGGAG----- 722
QY 1231 IleSerSerProAlaGluThrTrpIleThrCysGluAspArgIleArgValArgGly 1250
Db 723 -----GCCAGGCCCTGTGGAAGAGCGCGGCGCAATGTCACCATG---AGT 767
QY 1251 GlyAspThrGluCys-----SerGlyArg 1258
Db 768 GTGGATGCTGATGTGTGCTGCCCATGTCAGGGACCTTCTCAGGTACTTCTACTCCCGAAGG 827
QY 1259 ValGluIle-----TrpHisAlaGlySerTrpGlyThrValCys 1271
Db 828 ATTGACATCACCCCTGTCGTCAAGTCAAGTCTCCAC----- 863
QY 1272 AspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGly---CysGly 1290
Db 864 -----AAGCTGGCCTCTGCCTATGGGGCCGAGCAGCTGCGGGCTACTGCGCA 911
QY 1291 SerAlaLeuAlaLeu-----ArgAspAlaSerPheGlyGlnGlyThrGlyThrIle 1308
Db 912 AGCCTCTTTGCCATCTCTCTCCCGGAGACCCCTGCTTCCA-GATGCCCTGGACCTGTA 970
QY 1309 TrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLys 1328
Db 971 TGCCTA-----TGCAGTGGCCAC-----AGGGGAGCGC 997
QY 1329 ProTrpGlyClnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGln 1348
Db 998 CCTGCTGGA-----GAAGCTCTGCTACACTTCTGSCCTG 1033
QY 1349 SerLeuLysSerLeuAsnAlaSerSerGly 1358
Db 1034 GAACCTCGAGGCTTGACCGAGCGCGAGGC 1063

Search completed: May 12, 2003, 19:17:06
Job time : 236 secs